GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: October 18, 2002, 12:24:10 ; Search time 35 Seconds
(without alignments)
1720.057 Million cell updates/sec

Title: US-09-830-123-2
Perfect score: 2768
Sequence: 1 MAFGLSSLLQNSDLFTSDHA.....FGGRGFVPFVAGSPVEQSPR 542

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2768	100.0	542	22	AAB73251	Protein regulating
2	2127.5	76.9	555	21	AAB12786	Atliplex gmelini N
3	2124.5	76.8	553	22	AAB73253	Protein regulating
4	2103.5	76.0	540	22	AAB73252	Protein regulating
5	2055.5	74.3	535	21	AAB11832	Rice Na+/H+ antipo
6	2041	73.7	538	20	AAY40901	Arabidopsis thalia
7	1953.5	70.6	555	22	AAB73254	Protein regulating
8	1867	67.4	547	20	AAY40905	Arabidopsis thalia
9	1808	65.3	571	22	AAU02882	Arabidopsis thalia
10	1599.5	57.8	529	20	AAY40902	Arabidopsis thalia
11	620	22.4	649	22	ABB59364	Drosophila melanog
12	595	21.5	608	21	AAB29621	Cat flea HMT Na/H
13	575.5	20.8	727	22	ABB65651	Drosophila melanog
14	568	20.5	633	20	AAY44017	Saccharomyces cere
15	556	20.1	569	20	AAY44016	Schizosaccharomyce
16	537.5	19.4	509	22	AAB90555	Human secreted pro
17	537.5	19.4	526	22	AAB90637	Human secreted pro
18	483	17.4	518	22	ABB58504	Drosophila melanog
19	461	16.7	834	18	AAW11325	Human Na+/H+ excha
20	457.5	16.5	816	14	AAR44143	Rabbit sodium ion/
21	451	16.3	664	22	AAU02883	Human HsNHE-6 poly
22	445	16.1	325	20	AAY40906	Arabidopsis thalia
23	423	15.3	605	22	AAU02884	Yeast ScNHX1 polyp
24	354	12.8	273	22	AAU01607	Gene #21 human sec
25	348	12.6	275	22	AAU01608	Human secreted pro
26	310.5	11.2	315	22	AAM24062	Human EST encoded
27	299	10.8	339	22	AAB90590	Human secreted pro
28	299	10.8	339	22	AAB90591	Human secreted pro
29	290	10.5	359	22	AAU01581	Human secreted pro
30	290	10.5	990	22	ABG29884	Novel human diagno
31	289	10.4	232	20	AAY40907	Arabidopsis thalia
32	269.5	9.7	511	22	ABG10934	Novel human diagno
33	256.5	9.3	186	20	AAY40903	Arabidopsis thalia
34	255	9.2	707	22	AAM23715	Human EST encoded
35	255	9.2	707	22	AAM23744	Human EST encoded
36	251.5	9.1	684	22	AAU35119	Enterococcus faeca
37	236	8.5	557	22	ABB52506	Escherichia coli p
38	226.5	8.2	144	22	ABG10933	Novel human diagno
39	202	7.3	443	22	AAB96555	Putative P. abyssi
40	197.5	7.1	686	22	AAU34281	Staphylococcus aur
41	197.5	7.1	692	22	AAU37141	Staphylococcus aur
42	172.5	6.2	128	22	AAM24336	Drosophila EST enc
43	159.5	5.8	511	22	AAB76843	Corynebacterium gl
44	159.5	5.8	516	22	AAG91332	C glutamicum prote
45	151.5	5.5	178	20	AAY36089	Extended human sec

ALIGNMENTS

RESULT 1
AAB73251
ID AAB73251 standard; Protein; 542 AA.
XX
AC AAB73251;
XX
DT 14-MAY-2001 (first entry)
XX
DE Protein regulating the pH of vacuoles.
XX
KW Vacuole pH regulation; morning glory; flower colour.
XX
OS Ipomoea nil.
XX
PN WO200114560-A1.
XX
PD 01-MAR-2001.
XX
PF 24-AUG-2000; 2000WO-JP05722.
XX
PR 24-AUG-1999; 99JP-0236800.
XX
PA (SUNR) SUNTORY LTD.
XX
PI Iida S, Tanaka S, Inagaki Y;
XX
DR WPI; 2001-191648/19.
DR N-PSDB; AAF75752.
XX
PT Morning glory-originated gene encoding a protein with pH regulation
PT activity in vacuoles, useful in controlling flower color to give new
PT breeds of colorful plants for cut flowers, particularly applicable in
PT horticulture -
XX

```
PS   Claim 2; Page 30-33; 68pp; Japanese.
XX
CC   The present sequence is a protein from Morning Glory, which has vacuolar
CC   pH regulatory activities. The protein enables flower colour to be
CC   controlled via regulation of the vacuolar pH, colours can range from blue
CC   to red in colour spectrum. The protein is useful in controlling flower
CC   colour to give new breeds of colourful plants for cut flowers,
CC   particularly applicable in horticulture.
XX
SQ   Sequence   542 AA;

  Query Match            100.0%;  Score 2768;  DB 22;  Length 542;
  Best Local Similarity  100.0%;  Pred. No. 4.8e-270;
  Matches  542;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MAFGLSSLLQNSDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MAFGLSSLLQNSDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLC 60

Qy      61 TGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGT 120

Qy     121 LISCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVF 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LISCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVF 180

Qy     181 GEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIK 240

Qy     241 KLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 KLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVT 300

Qy     301 TRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVF 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVF 360

Qy     361 PLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMIT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 PLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMIT 420

Qy     421 STVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDS 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 STVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDS 480

Qy     481 ESDMITGPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVEQS 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ESDMITGPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVEQS 540

Qy     541 PR 542
           ||
Db     541 PR 542


RESULT 2
AAB12786
ID   AAB12786 standard; Protein; 555 AA.
XX
AC   AAB12786;
XX
DT   23-NOV-2000  (first entry)
XX
DE   Atliplex gmelini Na+ and H+ antiporter protein.
XX
KW   Atliplex gmelini; Na plus and H plus antiporter protein;
KW   Na+ and H+ antiporter protein; transformed plant; high salt tolerance.
XX
OS   Atliplex gmelini.
XX
PN   JP2000157287-A.
XX
PD   13-JUN-2000.
XX
PF   16-SEP-1999;   99JP-0261606.
XX
PR   24-SEP-1998;   98JP-0269504.
XX
PA   (SHOK-) SHOKUBUTSU KOGAKU KK.
XX
DR   WPI; 2000-468209/41.
DR   N-PSDB; AAA72926.
XX
PT   An Na+ and H+ plus antiporter protein and a gene encoding it  -
XX
PS   Disclosure; Page 10-12; 16pp; Japanese.
XX
CC   The present sequence represents an Na+ and H+ antiporter protein
CC   isolated from Atliplex gmelini. The Na+ and H+ antiporter protein and
CC   gene encoding it are useful for the preparation of transformed plants
CC   with high salt tolerance, e.g. for growth in arid land.
XX
SQ   Sequence   555 AA;

  Query Match            76.9%;  Score 2127.5;  DB 21;  Length 555;
  Best Local Similarity  76.6%;  Pred. No. 1.8e-205;
  Matches  418;  Conservative   50;  Mismatches   65;  Indels   13;  Gaps    3;

Qy       5 LSSLLQN--SDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLCTG 62
           |||||      | |||||||||||||||||| |||:|||||||||:|||||||:||| ||
Db       5 LSSLLSGKMDALTTSDHASVVSMNLFVALLCGCIVIGHLLEENRWMNESITALLIGLATG 64

Qy      63 VVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLI 122
           |||||:|||||||||||||||||||||||||||||||||||| ||:||:||||:|||:
Db      65 VVILLISGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFRNFITIVLFGAVGTLV 124

Qy     123 SCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVFGE 182
           | :||| ||: ||| ||| |:  |||||||||||||||||||:|||||||||||||||
Db     125 SFTIISLGALSIFKKLDIGTLELADYLAIGAIFAATDSVCTLQVLNQDETPLLYSLVFGE 184

Qy     183 GVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKL 242
           ||||||||||||||||||:|  | :| | |:|||||||::|| ||   ||| |||||||
Db     185 GVVNDATSVVLFNAIQSFDLTRIDHRIALQFMGNFLYLFIASTILGAFTGLLSAYIIKKL 244

Qy     243 YFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTR 302
           |||||||||||||||||:||||::||||||||||||||||||||||||||||||||||:
Db     245 YFGRHSTDREVALMMLMAYLSYMLAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTK 304

Qy     303 HSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPL 362
           |:|||||||| |:|||||||||||||:|| :| |:|||||||:||::|||||||||
Db     305 HAFATLSFVAEVFLFLYVGMDALDIEKWRFVSDSPGISVAVSSILLGLVMVGRAAFVFPL 364

Qy     363 SFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITST 422
           |:| | |||: |:|::| |||:||||||||||||:|||||:|| |||| |  ||||||||
Db     365 SWLMNFAKKSQSEKVTFNQQIVIWWAGLMRGAVSMALAYNQFTRSGHTQLRGNAIMITST 424

Qy     423 VTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSES 482
           ::||||||:||||:|||||  |||  |       |  | |:  |||  :::|||:   | |
Db     425 ISVVLFSTMVFGLLTKPLIMFLLPQPKHFT---SCSTVSDVGSPKSYSLPLLEGNQDYEV 481

Qy     483 DMITGPE--------VARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAG 534
           |:  |           : ||::|||||  |||||| |||||||||||||||||||||| |
Db     482 DVGNGNHEDTTEPRTIVRPSSLRMLLNAPTHTVHHYWRKFDDSFMRPVFGGRGFVPFVPG 541

Qy     535 SPVEQS 540
           || |||
Db     542 SPTEQS 547


RESULT 3
AAB73253
ID   AAB73253 standard; Protein; 553 AA.
XX
AC   AAB73253;
XX
```

```
DT   14-MAY-2001  (first entry)
XX
DE   Protein regulating the pH of vacuoles.
XX
KW   Vacuole pH regulation; flower colour.
XX
OS   Nierembergia hybrida.
XX
PN   WO200114560-A1.
XX
PD   01-MAR-2001.
XX
PF   24-AUG-2000; 2000WO-JP05722.
XX
PR   24-AUG-1999;   99JP-0236800.
XX
PA   (SUNR ) SUNTORY LTD.
XX
PI   Iida S,  Tanaka S,  Inagaki Y;
XX
DR   WPI; 2001-191648/19.
DR   N-PSDB; AAF75765.
XX
PT   Morning glory-originated gene encoding a protein with pH regulation
PT   activity in vacuoles, useful in controlling flower colo to give new
PT   breeds of colorful plants for cut flowers, particularly applicable in
PT   horticulture -
XX
PS   Example 7; Page 50-52; 68pp; Japanese.
XX
CC   The present sequence is a protein, which has vacuolar pH regulatory
CC   activities. The protein enables flower colour to be controlled via
CC   regulation of the vacuolar pH, colours can range from blue to red in
CC   colour spectrum. The protein is useful in controlling flower colour to
CC   give new breeds of colourful plants for cut flowers, particularly
CC   applicable in horticulture.
XX
SQ   Sequence   553 AA;

  Query Match            76.8%;  Score 2124.5;  DB 22;  Length 553;
  Best Local Similarity  77.0%;  Pred. No. 3.7e-205;
  Matches  415;  Conservative   51;  Mismatches   66;  Indels    7;  Gaps    3;

Qy       1 MAFGLSSLL-QNSDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGL 59
           |||   :|| : ::| |||| |||||:||||||:|||||:|||||||||:|||||||:||
Db       1 MAFDFGTLLGKMNNLTTSDHQSVVSVNLFVALICACIVIGHLLEENRWMNESITALVIGS 60

Qy      60 CTGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIG 119
           ||||:|||:||||:||:|||||||||||||||||||||||||| || || |||||||:|
Db      61 CTGVIILLISGGKNSHILVFSEDLFFIYLLPPIIFNAGFQVKKKSFFRNFSTIMLFGAVG 120

Qy     120 TLISCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLV 179
           |||| ||| ||: ||| :||  |: |||||||||||||||||||||:|:|||||||||
Db     121 TLISFIIISAGAIGIFKKMDIGHLEIGDYLAIGAIFAATDSVCTLQVLNQEETPLLYSLV 180

Qy     180 FGEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYII 239
           |||||||||||||||||:|:||::       |  |||||||| ||||||| :||| |:||
Db     181 FGEGVVNDATSVVLFNAVQNFDLSHISTGKALQLIGNFLYLFASSTFLGVAVGLLSAFII 240

Qy     240 KKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRV 299
           |||||||||||||||:|:||:||||::|||||||||||||||||||||||||||||||||
Db     241 KKLYFGRHSTDREVAIMILMAYLSYMLAELFYLSGILTVFFCGIVMSHYTWHNVTESSRV 300

Qy     300 TTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFV 359
           ||:|:|||||||:|| |||||||||||||||||| :| | |: |||||:||:|||| |||
Db     301 TTKHTFATLSFIAEIFIFLYVGMDALDIEKWKFVSDSPGTSIKVSSILLGLVLVGRGAFV 360

Qy     360 FPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMI 419
           |||||||| ||| |||| ||: |||||||||||:|||||:||  ||| |  ||||| 
Db     361 FPLSFLSNLTKKNPEDKISFNQQVTIWWAGLMRGAVSMALAYNQFTRGGHTQLRANAIMI 420

Qy     420 TSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPD 479
           |||:||||||||||||||||| |||| | :      | :||| :|| | |||||:  |
Db     421 TSTITVVLFSTVVFGLMTKPLILLLLPSQKHL----IRMISSEPMTPKSFIVPLLDSTQD 476

Qy     480 SESDMITGPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVE 538
           ||:|:  |  | || :||||| ||:|||| ||||||::||||||||||||| ||| |
Db     477 SEADL--GRHVPRPHSLRMLLSTPSHTVHYYWRKFDNAFMRPVFGGRGFVPFVPGSPTE 533


RESULT 4
AAB73252
ID   AAB73252 standard; Protein; 540 AA.
XX
AC   AAB73252;
XX
DT   14-MAY-2001  (first entry)
XX
DE   Protein regulating the pH of vacuoles.
XX
KW   Vacuole pH regulation; flower colour.
XX
OS   Petunia hybrida.
XX
PN   WO200114560-A1.
XX
PD   01-MAR-2001.
XX
PF   24-AUG-2000; 2000WO-JP05722.
XX
PR   24-AUG-1999;   99JP-0236800.
XX
PA   (SUNR ) SUNTORY LTD.
XX
PI   Iida S,  Tanaka S,  Inagaki Y;
XX
DR   WPI; 2001-191648/19.
DR   N-PSDB; AAF75764.
XX
PT   Morning glory-originated gene encoding a protein with pH regulation
PT   activity in vacuoles, useful in controlling flower color to give new
PT   breeds of colorful plants for cut flowers, particularly applicable in
PT   horticulture -
XX
PS   Example 7; Page 42-45; 68pp; Japanese.
XX
CC   The present sequence is a protein, which has vacuolar pH regulatory
CC   activities. The protein enables flower colour to be controlled via
CC   regulation of the vacuolar pH, colours can range from blue to red in
CC   colour spectrum. The protein is useful in controlling flower colour to
CC   give new breeds of colourful plants for cut flowers, particularly
CC   applicable in horticulture.
XX
SQ   Sequence   540 AA;

  Query Match            76.0%;  Score 2103.5;  DB 22;  Length 540;
  Best Local Similarity  76.4%;  Pred. No. 4.7e-203;
  Matches  412;  Conservative   51;  Mismatches   69;  Indels    7;  Gaps    3;

Qy       1 MAFGLSSLLQNSD-LFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGL 59
           |||   :|| | | | |||| ||||:||||||:||||||:|||||||||:|||||||:||
Db       1 MAFDFGTLLGNVDRLSTSDHQSVVSINLFVALICACIVIGHLLEENRWMNESITALVIGS 60

Qy      60 CTGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIG 119
           |||:||||:||||:||:|||||||||||||||||||||||||| || || |||||||:|
Db      61 CTGIVILLISGGKNSHILVFSEDLFFIYLLPPIIFNAGFQVKKKSFFRNFSTIMLFGALG 120

Qy     120 TLISCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLV 179
           |||| ||| ||: ||| ::|  |: |||||||||:||||||||||:||||||||||
Db     121 TLISFIIISLGAIGIFKKMNIGSLEIGDYLAIGAIFSATDSVCTLQVLNQDETPLLYSLV 180

Qy     180 FGEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYII 239
           |||||||||||||||||:||::  |    :  :|||||| ||| |||  ||| ||||
Db     181 FGEGVVNDATSVVLFNAIQNFDLSHIDTGKAMELVGNFLYLFASSTALGVAAGLLSAYII 240

Qy     240 KKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRV 299
```

```
Db    241 KKLYFGRHSTDREVAIMILMAYLSYMLAELFYLSAILTVFFSGIVMSHYTWHNVTESSRV 300

Qy    300 TTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFV 359

Db    301 TTKHTFATLSFIAEIFIFLYVGMDALDIEKWKFVSDSPGISVQVSSILLGLVLVGRAAFV 360

Qy    360 FPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMI 419

Db    361 FPLSFLSNLTKKTPEAKISFNQQVTIWWAGLMRGAVSMALAYNQFTRGGHTQLRANAIMI 420

Qy    420 TSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPD 479

Db    421 TSTITVVLFSTVVFGLMTKPLIRILLPSHKHL----SRMISSEPTTPKSFIVPLLDSTQD 476

Qy    480 SESDMITGPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVE 538

Db    477 SEADL--ERHVPRPHSLRMLLSTPSHTVHYYWRKFDNAFMRPVFGGRGFVPFAPGSPTD 533


RESULT 5
AAB11832
ID   AAB11832 standard; Protein; 535 AA.
XX
AC   AAB11832;
XX
DT   07-NOV-2000  (first entry)
XX
DE   Rice Na+/H+ antiporter, OsNHX1.
XX
KW   OsNHX1; Na+/H+ antiporter; sodium/proton antiporter; countertransporter;
KW   active transport; rice; transgenic plant; salt-tolerance.
XX
OS   Oryza sativa.
XX
PN   WO200037644-A1.
XX
PD   29-JUN-2000.
XX
PF   22-DEC-1999;   99WO-JP07224.
XX
PR   22-DEC-1998;   98JP-0365604.
XX
PA   (NORQ ) JAPAN MIN AGRIC FORESTRY & FISHERIES.
XX
PI   Fukuda A,  Tanaka Y;
XX
DR   WPI; 2000-442672/38.
DR   N-PSDB; AAA61876.
XX
PT   DNA encoding a sodium ion and proton counter-transporter protein of
PT   rice origin for production of salt tolerant rice transformants  -
XX
PS   Claim 1a; Fig 1; 43pp; Japanese.
XX
CC   This sequence represents a novel rice Na+/H+ antiporter
CC   (countertransporter), OsNHX1. The invention relates to OsNHX1 and
CC   nucleic acids which encode it; vectors, host cells and transgenic plants
CC   containing OsNHX1 nucleic acids; recombinant expression of OsNHX1; and
CC   antibodies which recognise OsNHX1. OsNHX1 nucleic acids are useful in the
CC   production of salt tolerant transgenic plants.
XX
SQ   Sequence   535 AA;

  Query Match             74.3%;  Score 2055.5;  DB 21;  Length 535;
  Best Local Similarity   76.4%;  Pred. No. 3.2e-198;
  Matches  401;  Conservative   49;  Mismatches   66;  Indels    9;  Gaps    3;

Qy     16 TSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLCTGVVILLLSGGKSSH 75

Db     16 TSDYASVVSINLFVALLCACIVLGHLLEENRWVNESITALIIGLCTGVVILLMTKGKSSH 75

Qy     76 LLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIISFGAVKIF 135

Db     76 LFVFSEDLFFIYLLPPIIFNAGFQVKKKQFFRNFMTITLFGAVGTMISFFTISIAAIAIF 135

Qy    136 KHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVFGEGVVNDATSVVLFN 195

Db    136 SRMNIGTLDVGDFLAIGAIFSATDSVCTLQVLNQDETPFLYSLVFGEGVVNDATSIVLFN 195

Qy    196 AIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDREVAL 255

Db    196 ALQNFDLVHIDAAVVLKFLGNFFYLFLSSTFLGVFAGLLSAYIIKKLYIGRHSTDREVAL 255

Qy    256 MMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETF 315

Db    256 MMLMAYLSYMLAELLDLSGILTVFFCGIVMSHYTWHNVTESSRVTTKHAFATLSFIAETF 315

Qy    316 IFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKKNSSD 375

Db    316 LFLYVGMDALDIEKWEFASDRPGKSIGISSILLGLVLIGRAAFVFPLSFLSNLTKKAPNE 375

Qy    376 KISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTVVFGL 435

Db    376 KITWRQQVVIWWAGLMRGAVSIALAYNKFTRSGHTQLHGNAIMITSTITVVLFSTMVFGM 435

Qy    436 MTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMITGPEVARPTA 495

Db    436 MTKPLIRLLLP-----ASGHP--VTSEPSSPKSLHSPLLTSMQGSDLESTT--NIVRPSS 486

Qy    496 LRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVEQS 540

Db    487 LRMLLTKPTHTVHYYWRKFDDALMRPMFGGRGFVPFSPGSPTEQS 531


RESULT 6
AAY40901
ID   AAY40901 standard; Protein; 538 AA.
XX
AC   AAY40901;
XX
DT   18-JAN-2000  (first entry)
XX
DE   Arabidopsis thaliana Na/H transporter AtNHX1.
XX
KW   Sodium; proton; antiport; transporter; salt tolerance; salt management;
KW   transgenic plant; survival; soil; farming; accumulation; irrigation;
KW   crop.
XX
OS   Arabidopsis thaliana.
XX
PN   WO9947679-A2.
XX
PD   23-SEP-1999.
XX
PF   18-MAR-1999;   99WO-CA00219.
XX
PR   18-MAR-1998;   98US-0078474.
PR   15-JAN-1999;   99US-0116111.
XX
PA   (BLUM/) BLUMWALD E.
PA   (APSE/) APSE M.
PA   (SNED/) SNEDDEN W.
PA   (AHAR/) AHARON G.
XX
PI   Blumwald E,  Apse M,  Snedden W,  Aharon G;
XX
DR   WPI; 1999-571840/48.
DR   N-PSDB; AAZ22591.
XX
PT   Nucleic acid molecules encoding sodium/proton transport polypeptides,
PT   useful in genetic engineering salt tolerance in crop plants  -
XX
PS   Claim 36; Fig 1A; 93pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule encoding
```

```
CC   a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment
CC   and capable of increasing salt tolerance in a cell. This sequence
CC   corresponds to the AtNHX1 transporter from Arabidopsis thaliana.
CC   The Na/H transporter polypeptides provide a means of intracellular
CC   salt management, particularly in plants. The sequences are useful for
CC   producing transgenic plants that are capable of surviving in soil with
CC   high salt levels that would normally inhibit growth of the crop species.
CC   This would be useful in farming land in areas that are generally
CC   considered unproductive through salt accumulation and poor irrigation,
CC   e.g. in India, Australia, and prairies in USA or Canada. Commercial
CC   crops, such as potatoes, tomatoes, brassica, cotton, sunflower,
CC   strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley,
CC   atriplex, sorghum, alfalfa, salicornia and others would benefit from
CC   increased salt tolerance.
XX
SQ   Sequence   538 AA;

  Query Match             73.7%;  Score 2041;  DB 20;  Length 538;
  Best Local Similarity   73.8%;  Pred. No. 9.3e-197;
  Matches  399;  Conservative   57;  Mismatches   71;  Indels   14;  Gaps    4;

Qy       5 LSSLLQN-SDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLCTGV 63
Db       2 LDSLVSKLPSLSTSDHASVVALNLFVALLCACIVLGHLLEENRWMNESITALLIGLGTGV 61

Qy      64 VILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLIS 123
Db      62 TILLISKGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFRNFVTIMLFGAVGTIIS 121

Qy     124 CSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVFGEG 183
Db     122 CTIISLGVTQFFKKLDIGTFDLGDYLAIGAIFAATDSVCTLQVLNQDETPLLYSLVFGEG 181

Qy     184 VVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKLY 243
Db     182 VVNDATSVVVFNAIQSFDLTHLNHEAAFHLLGNFLYLFLLSTLLGAATGLISAYVIKKLY 241

Qy     244 FGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRH 303
Db     242 FGRHSTDREVALMMLMAYLSYMLAELFDLSGILTVFFCGIVMSHYTWHNVTESSRITTKH 301

Qy     304 SFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLS 363
Db     302 TFATLSFLAETFIFLYVGMDALDIDKWRSVSDTPGTSIAVSSILMGLVMVGRAAFVFPLS 361

Qy     364 FLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTV 423
Db     362 FLSNLAKKNQSEKINFNMQVVIWWSGLMRGAVSMALAYNKFTRAGHTDVRGNAIMITSTI 421

Qy     424 TVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTT---SEPSSPKHFTVPLLDNQPDS 480
Db     422 TVCLFSTVVFGMLTKPLISYLLP--------HQNATTSMLSDDNTPKSIHIPLLDQ--DS 471

Qy     481 ESDMITGPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVEQS 540
Db     472 FIEPSGNHNVPRPDSIRGFLTRPTRTVHYYWRQFDDSFMRPVFGGRGFVPFVPGSPTERN 531

Qy     541 P 541
Db     532 P 532


RESULT 7
AAB73254
ID   AAB73254 standard; Protein; 555 AA.
XX
AC   AAB73254;
XX
DT   14-MAY-2001  (first entry)
XX
DE   Protein regulating the pH of vacuoles.
XX
KW   Vacuole pH regulation; flower colour.
XX
OS   Torenia hybrida.
XX
PN   WO200114560-A1.
XX
PD   01-MAR-2001.
XX
PF   24-AUG-2000; 2000WO-JP05722.
XX
PR   24-AUG-1999;   99JP-0236800.
XX
PA   (SUNR ) SUNTORY LTD.
XX
PI   Iida S,  Tanaka S,  Inagaki Y;
XX
DR   WPI; 2001-191648/19.
DR   N-PSDB; AAF75766.
XX
PT   Morning glory-originated gene encoding a protein with pH regulation
PT   activity in vacuoles, useful in controlling flower color to give new
PT   breeds of colorful plants for cut flowers, particularly applicable in
PT   horticulture -
XX
PS   Example 7; Page 57-60; 68pp; Japanese.
XX
CC   The present sequence is a protein, which has vacuolar pH regulatory
CC   activities. The protein enables flower colour to be controlled via
CC   regulation of the vacuolar pH, colours can range from blue to red in
CC   colour spectrum. The protein is useful in controlling flower colour to
CC   give new breeds of colourful plants for cut flowers, particularly
CC   applicable in horticulture.
XX
SQ   Sequence   555 AA;

  Query Match             70.6%;  Score 1953.5;  DB 22;  Length 555;
  Best Local Similarity   70.5%;  Pred. No. 6.4e-188;
  Matches  389;  Conservative   56;  Mismatches   88;  Indels   19;  Gaps    4;

Qy       4 GLSSLL-----QNSDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIG 58
Db       2 GFESVIKLAASETDNLWSSGHGSVVAITLFVTLLCTCIVIGHLLEENRWMNESIIALIIG 61

Qy      59 LCTGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAI 118
Db      62 LATGVIILLISGGKSSHLLVFSEDLFFIYALPPIIFNAGFQVKKKSFFRNFATIMMFGAV 121

Qy     119 GTLISCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSL 178
Db     122 GTLISFIIISLGTIAFFPKMNMR-LGVGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSL 180

Qy     179 VFGEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYI 238
Db     181 VFGEGVVNDATSVVLFNAVQNFDLPHMSTAKAFELVGNFFYLFATSTVLGVLTGLLSAYI 240

Qy     239 IKKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSR 298
Db     241 IKKLYFGRHSTDREVAIMILMAYLSYMLAELFDLSGILTVFFCGIVMSHYTWHNVTENSR 300

Qy     299 VTTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAF 358
Db     301 VTTKHTFATLSFVAEIFIFLYVGMDALDIEKWRFVSGSMTTSAAVSATLLGLVLLSRAAF 360

Qy     359 VFPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIM 418
Db     361 VFPLSFLSNLAKKSPLEKISLRQQIIIWWAGLMRGAVSMALAYKQFTREGLTVERENAIF 420

Qy     419 ITSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQP 478
Db     421 ITSTITIVLFSTVVFGLMTKPLINLLIPSPKLNRS-----VSSEPLTPNSITIPLLGESQ 475

Qy     479 DSESDMIT--------GPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVP 530
Db     476 DSVAELFSIRGQTSQGGEPVARPSSLRMLLTKPTHTVHYYWRKFDNAFMRPVFGGRGFVP 535
```

```
Qy     531 FVAGSPVEQSPR 542
           :| ||| |:| |
Db     536 YVPGSPTERSVR 547


RESULT 8
AAY40905
ID   AAY40905 standard; Protein; 547 AA.
XX
AC   AAY40905;
XX
DT   18-JAN-2000  (first entry)
XX
DE   Arabidopsis thaliana Na/H transporter.
XX
KW   Sodium; proton; antiport; transporter; salt tolerance; salt management;
KW   transgenic plant; survival; soil; farming; accumulation; irrigation;
KW   crop.
XX
OS   Arabidopsis thaliana.
XX
PN   WO9947679-A2.
XX
PD   23-SEP-1999.
XX
PF   18-MAR-1999;   99WO-CA00219.
XX
PR   18-MAR-1998;   98US-0078474.
PR   15-JAN-1999;   99US-0116111.
XX
PA   (BLUM/) BLUMWALD E.
PA   (APSE/) APSE M.
PA   (SNED/) SNEDDEN W.
PA   (AHAR/) AHARON G.
XX
PI   Blumwald E,  Apse M,  Snedden W,  Aharon G;
XX
DR   WPI; 1999-571840/48.
DR   N-PSDB; AAZ22595.
XX
PT   Nucleic acid molecules encoding sodium/proton transport polypeptides,
PT   useful in genetic engineering salt tolerance in crop plants -
XX
PS   Claim 50; Fig 5A-B; 93pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule encoding
CC   a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment
CC   and capable of increasing salt tolerance in a cell.  This sequence
CC   corresponds to a transporter from Arabidopsis thaliana.
CC   The Na/H transporter polypeptides provide a means of intracellular
CC   salt management, particularly in plants. The sequences are useful for
CC   producing transgenic plants that are capable of surviving in soil with
CC   high salt levels that would normally inhibit growth of the crop species.
CC   This would be useful in farming land in areas that are generally
CC   considered unproductive through salt accumulation and poor irrigation,
CC   e.g. in India, Australia, and prairies in USA or Canada. Commercial
CC   crops, such as potatoes, tomatoes, brassica, cotton, sunflower,
CC   strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley,
CC   atriplex, sorghum, alfalfa, salicornia and others would benefit from
CC   increased salt tolerance.
XX
SQ   Sequence   547 AA;

  Query Match             67.4%;  Score 1867;  DB 20;  Length 547;
  Best Local Similarity   73.2%;  Pred. No. 3.3e-179;
  Matches  369;  Conservative   54;  Mismatches   67;  Indels   14;  Gaps    4;

Qy       5 LSSLLQN-SDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLCTGV 63
           | ||:     | ||||||||::||||||||||||||||||||:|||||||:||| |||
Db       2 LDSLVSKLPSLSTSDHASVVALNLFVALLCACIVLGHLLEENRWMNESITALLIGLGTGV 61

Qy      64 VILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLIS 123
                                                                  |||:| ||||||||||||||||||||||||||||||||||| ||:|||||||:||:||
Db      62 TILLISKGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFRNFVTIMLFGAVGTIIS 121

Qy     124 CSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVFGEG 183
           |:||| |  : || |||   | ||||||||||||||||||||:|||||||||||||||
Db     122 CTIISLGVTQFFKKLDIGTFDLGDYLAIGAIFAATDSVCTLQVLNQDETPLLYSLVFGEG 181

Qy     184 VVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKLY 243
           |||||||||:|||||||:|  : :   | :|||||||| || ||   ||: ||:|||||
Db     182 VVNDATSVVVFNAIQSFDLTHLNHEAAFHLLGNFLYLFLLSTLLGAATGLISAYVIKKLY 241

Qy     244 FGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRH 303
           |||||||||||||||:||||::|||| ||||||||||||||||||||||||:||:|
Db     242 FGRHSTDREVALMMLMAYLSYMLAELFDLSGILTVFFCGIVMSHYTWHNVTESSRITTKH 301

Qy     304 SFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLS 363
           :||||||:|||||||||||||||||:||: | :: | |:||||||:||::|||||||||
Db     302 TFATLSFLAETFIFLYVGMDALDIDKWRSVSDTPGTSIAVSSILMGLVMVGRAAFVFPLS 361

Qy     364 FLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTV 423
           |||||||| |:||:|  |::|||:||||||||:||||||| :||| :  ||||||||:
Db     362 FLSNLAKKNQSEKINFNMQVVIWWSGLMRGAVSMALAYNKFTRAGHTDVRGNAIMITSTI 421

Qy     424 TVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTT---SEPSSPKHFTVPLLDNQPDS 480
           || |||||||::|||||: |||         | : ||   |: ::||   :||||   ||
Db     422 TVCLFSTVVFGMLTKPLISYLLP--------HQNATTSMLSDDNTPKSIHIPLLDQ--DS 471

Qy     481 ESDMITGPEVARPTALRMLLRTPT 504
             :       | || ::|  |  ||
Db     472 FIEPSGNHNVPRPDSIRGFLTRPT 495


RESULT 9
AAU02882
ID   AAU02882 standard; Protein; 571 AA.
XX
AC   AAU02882;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Arabidopsis thaliana AtNHX1 polypeptide.
XX
KW   AtNHX1; tonoplast pyrophosphatase hydrogen ion translocating pump; AVP1;
KW   vacuolar pyrophosphatase; salt tolerance; drought tolerance; cation;
KW   freeze tolerance; 35S promoter; sodium; calcium; manganese; lead;
KW   saline soil; thale-cress.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200133945-A1.
XX
PD   17-MAY-2001.
XX
PF   10-NOV-2000; 2000WO-US30955.
XX
PR   10-NOV-1999;   99US-0164808.
PR   18-AUG-2000; 2000US-0226223.
PR   22-AUG-2000; 2000US-0644039.
XX
PA   (UYCO-) UNIV CONNECTICUT.
PA   (WHED ) WHITEHEAD INST.
XX
PI   Gaiola RA;
XX
DR   WPI; 2001-328862/34.
XX
PT   Transgenic plants with salt tolerance, drought and freeze resistance
PT   and increased yield and flower size comprise exogenous H+ genes or
PT   increased vacuolar pyrophosphatase expression -
XX
PS   Example 1; Fig 6; 68pp; English.
XX
CC   The sequence represents an Arabidopsis thaliana AtNHX1 polypeptide which
```

```
CC   is a tonoplast pyrophosphatase hydrogen ion (H+) translocating pump.
CC   Transgenic plants can be made by transforming plant cells with exogenous
CC   tonoplast pyrophosphate driven H+ pump genes and an exogenous nucleic
CC   acid encoding a protein, such as AVP1, which alters expression of
CC   vacuolar pyrophosphatase. Salt tolerance may be introduced into a plant
CC   via transformation of the cells to induce upregulation of vacuolar
CC   phosphasphatase expression. Drought and/or freeze tolerance may also be
CC   introduced through transformation with DNA encoding a vacuolar H+
CC   translocating pump linked to a promoter such as the 35S promoter. These
CC   processes are useful for bioremediating soil and removing cations such as
CC   sodium, calcium, manganese and lead from soil or water which can
CC   support plant growth. Plants which grow in saline soil can be produced
CC   and yield and flower size of plants can be increased.
XX
SQ   Sequence   571 AA;

  Query Match             65.3%;  Score 1808;  DB 22;  Length 571;
  Best Local Similarity   68.6%;  Pred. No. 3.1e-173;
  Matches  371;  Conservative   57;  Mismatches   99;  Indels   14;  Gaps    4;

Qy       5 LSSLLQN-SDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLCTGV 63
Db       2 LDSLVSKLPSLSTSDHASVVALNLPVALLCACIVLGHLLSSNRWMNKSITALLIGLGTGV 61

Qy      64 VILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLIS 123
Db      62 TILLISKGKSSHLLVFSEDLFPIYLLPPIIPNACFQVKKKQFFHNFVTINLPCAVGTIIS 121

Qy     124 CSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVFGEG 183
Db     122 CTIISLCVTQFPKKLDIGTFDLCDYLAICAIFAATDSVCTLQVLNQDETPLLYSLVFCEC 181

Qy     184 VVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKLY 243
Db     182 VVNDATSVVVFNAIQSFDLTHLNHRAAFHLLCNFLYLFLLSTLLGAATCLISAYVIKKLY 241

Qy     244 FGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRH 303
Db     242 FCRHSTDRRVALMMLMAYLSYMLARLFDLSCILTVPFCGIVMSHYTWHNVTESSRITTKH 301

Qy     304 SFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLS 363
Db     302 TFATLSFLARTFIFLYVCMDALDIDKWRSVSDTPGTSIAVSSILMGLVMVGRAAFVFPLS 361

Qy     364 FLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTV 423
Db     362 FLSNLAKKNQSRKINFNMQVVIWWSGLMRCAVSMALAYNKFTRAGHTDVRGNAIMITSTI 421

Qy     424 TVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTT---SEPSSPKHFTVPLLDNQPDS 480
Db     422 TVCLFSTVVFGMLTKPLISYLLP--------HQNATTSMLSDDNTPKSIHIPLLDQ--DS 471

Qy     481 ESDMITGPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVEQS 540
Db     472 FIEPSCNHNVPRPDSIRGFLTRPTRTVHYYWRQFDDSFMRPVFCCRGFVPFVPGSPTERN 531

Qy     541 P 541
Db     532 P 532


RESULT 10
AAY40902
ID   AAY40902 standard; Protein; 529 AA.
XX
AC   AAY40902;
XX
DT   18-JAN-2000  (first entry)
XX
DE   Arabidopsis thaliana Na/H transporter AtNHX2.
XX
KW   Sodium; proton; antiport; transporter; salt tolerance; salt management;
KW   transgenic plant; survival; soil; farming; accumulation; irrigation;
KW   crop.
XX
OS   Arabidopsis thaliana.
XX
PN   WO9947679-A2.
XX
PD   23-SEP-1999.
XX
PF   18-MAR-1999;   99WO-CA00219.
XX
PR   18-MAR-1998;   98US-0078474.
PR   15-JAN-1999;   99US-0116111.
XX
PA   (BLUM/) BLUMWALD E.
PA   (APSE/) APSE M.
PA   (SNED/) SNEDDEN W.
PA   (AHAR/) AHARON G.
XX
PI   Blumwald E,  Apse M,  Snedden W,  Aharon G;
XX
DR   WPI; 1999-571840/48.
DR   N-PSDB; AAZ22592.
XX
PT   Nucleic acid molecules encoding sodium/proton transport polypeptides,
PT   useful in genetic engineering salt tolerance in crop plants -
XX
PS   Claim 36; Fig 1B; 93pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule encoding
CC   a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment
CC   and capable of increasing salt tolerance in a cell.  This sequence
CC   corresponds to the AtNHX2 transporter from Arabidopsis thaliana.
CC   The Na/H transporter polypeptides provide a means of intracellular
CC   salt management, particularly in plants. The sequences are useful for
CC   producing transgenic plants that are capable of surviving in soil with
CC   high salt levels that would normally inhibit growth of the crop species.
CC   This would be useful in farming land in areas that are generally
CC   considered unproductive through salt accumulation and poor irrigation,
CC   e.g. in India, Australia, and prairies in USA or Canada. Commercial
CC   crops, such as potatoes, tomatoes, brassica, cotton, sunflower,
CC   strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley,
CC   atriplex, sorghum, alfalfa, salicornia and others would benefit from
CC   increased salt tolerance.
XX
SQ   Sequence   529 AA;

  Query Match             57.8%;  Score 1599.5;  DB 20;  Length 529;
  Best Local Similarity   58.7%;  Pred. No. 2.9e-152;
  Matches  310;  Conservative   87;  Mismatches  120;  Indels   11;  Gaps    5;

Qy       1 MAFGLSSLLQNSDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLC 60
Db       1 MSIGLTEFVTNK--LAAEHPQVIPISVFIAILCLCLVIGHLLEENRWVNESITAILVGAA 58

Qy      61 TGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGT 120
Db      59 SGTVILLISKGKSSHILVFDEELFFIYLLPPIIFNAGFQVKKKKFFHNFLTIMSFGVIGV 118

Qy     121 LISCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVF 180
Db     119 FISTVIISFGTWWLFPKLGFKGLSARDYLAIGTIFSSTDTVCTLQILHQDETPLLYSLVF 178

Qy     181 GEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIK 240
Db     179 GEGVVNDATSVVLFNAVQKIQFESLTGWTALQVFGNFLYLFSTSTLLGIGVGLITSFVLK 238

Qy     241 KLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVT 300
Db     239 TLYFGRHSTTRELAIMVLMAYLSYMLAELFSLSGILTVFFCGVLMSHYASYNVTESSRIT 298

Qy     301 TRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVF 360
Db     299 SRHVFAMLSFIAETFIFLYVGTDALDFTKWKTSSLSFGGTLGVSGVITALVLLGRAAFVF 358
```

```
Qy     361 PLSFLSNLAKKNS--SDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIM 418
           ||| |:|   :::  :: |:|: |:||||||||||||||||||: :|| || |    || |
Db     359 PLSVLTNFMNRHTERNESITFKHQVIIWWAGLMRGAVSIALAFKQFTYSGVTLDPVNAAM 418

Qy     419 ITSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKH-FTVPLLDNQ 477
           :|:|  ||||:|:||| :||||:| |||      :|:    | || |||   |:|||
Db     419 VTNTTIVVLFTTLVFGFLTKPLVNYLLPQDASHNTGNRGKRT-EPGSPKEDATLPLLSFD 477

Qy     478 PDSESDMITGPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGG 525
           : ::    :    :: :|:  | :|:||||||||::|||:|||
Db     478 ESASTNFNRARD-----SISLLMEQPVYTIHRYWRKFDDTYMRPIFGG 520


RESULT 11
ABB59364
ID   ABB59364 standard; Protein; 649 AA.
XX
AC   ABB59364;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 4884.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL03467.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 4884; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   649 AA;

  Query Match             22.4%;  Score 620;  DB 22;  Length 649;
  Best Local Similarity   33.3%;  Pred. No. 2.2e-53;
  Matches  159;  Conservative  100;  Mismatches  173;  Indels   46;  Gaps   14;

Qy       6 SSLLQNSDLFTS-----DHASVVSMNLFVALLCACIVLGHLLEEN--RWVNESITALIIG 58
           |:|: :  |  |     :| | :|:   : ::   |:| | : :   ::: |||  : :|
Db      80 STLVTDPPLIDSHAVEQEHNSSLSLFFVICVIMLGILLIHSMLQTGFQYLPESIVVVFLG 139

Qy      59 LCTGVVILLLSG--GKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFG 116
           |: : ::||  |      |||   ||: |||||| :|: : |  || |  :|::|
Db     140 AFIGLSLNVMSGQNGSWKREEVFSPMGFFLVLLPPIIFESGYNLHKGNFFQNIGSILVFA 199

Qy     117 AIGTLISCSIISFGAVKIFKHLDIDF-LDFGDYLAIGAIFAATDSVCTLQVL-SQDETPL 174
           || ||  :|  | : :    :: | | | :  | |:: :| | | |: :  : |  |:
Db     200 IFGTTISALVIGAG-IYLLGLGEVAFRLSFSESFAFGSLISAVDPVATVAIFHALDVDPI 258

Qy     175 LYSLVFGEGVVNDATSVVLFNAIQSFDMTSFDPKIG---LHFIGNFLYLFLSSTFLGVGI 231
           |  ||||| ::||| |:||  :|      : :   |      :  |  :| :|  :||
Db     259 LNMLVFGESILNDAISIVLTASITQSANVNAEASTGEAMFSALKTFCAMFFASAGIGVIF 318

Qy     232 GLLCAYIIKKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWH 291
           |: | ::| :   :| :  | |:|:: :|  |::|| :|||: : ||||||||||
Db     319 ALISALLLKHIDLRKHPS-LEFAMMLMFTYAPYVLAEGIHLSGIMAILFCGIVMSHYTHF 377

Qy     292 NVTESSRVTTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLI 351
           |::  :::| : :  ||:|:||| :| |:|:           |:   ||  : :|:  |
Db     378 NLSTVTQITMQQTMRTLAFIAETCVFAYLGLAIFSF------KHQVELSFVIWAIV--LC 429

Qy     352 LVGRAAFVFPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTS 411
           |:|||  :|||:|| |   |    ||: : | |:|::|| |||:| ||          |
Db     430 LIGRACNIFPLAFLVN---KFREHKINNKMQFIMWFSGL-RGAISYAL----------S 474

Qy     412 LHENA-------IMITSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSE 462
           || |         ::||:|: :|||:|:| |  | ||:  | |  |:   |       |
Db     475 LHLNLDSQEKRHVIITTTLIIVLFTTLVLGGSTMPLLKYLKPGKKRRARGSGRNAAEE 532


RESULT 12
AAB29621
ID   AAB29621 standard; Protein; 608 AA.
XX
AC   AAB29621;
XX
DT   19-FEB-2001  (first entry)
XX
DE   Cat flea HMT Na/H transporter, SEQ ID NO:1868.
XX
KW   Cat flea; hindgut and Malpighian tubule protein; HMT;
KW   flea infestation; vaccine; antiparasitic; therapeutic target;
KW   diagnosis; detection.
XX
OS   Ctenocephalides felis.
XX
PN   WO200061621-A2.
XX
PD   19-OCT-2000.
XX
PF   07-APR-2000; 2000WO-US09437.
XX
PR   09-APR-1999;   99US-0128704.
XX
PA   (HESK-) HESKA CORP.
XX
PI   Brandt KS,  Gaines PJ,  Stinchcomb DT,  Wisnewski N;
XX
DR   WPI; 2000-656323/63.
DR   N-PSDB; AAC95370, AAC95371, AAC95372, AAC95373.
XX
PT   Flea Malpighian tubule and head and nerve cord tissue derived nucleic
PT   acids useful for the prevention, diagnosis and treatment of flea
PT   infestations -
XX
PS   Claim 10; Page 891-893; 964pp; English.
XX
CC   The invention relates to novel cat flea (Ctenocephalides felis) nucleic
CC   acids which are expressed in hindgut and Malpighian tubule (HMT) tissue
CC   or head and nerve cord (HNC) tissue. The invention also relates to the
CC   encoded proteins. The invention additionally encompasses expression
CC   constructs, recombinant viruses and recombinant cells comprising the
CC   nucleic acids of the invention, recombinant production of the proteins,
CC   antibodies against the proteins, a method of identifying inhibitors of
```

```
CC   the proteins, and compositions comprising the inhibitors for
CC   administration to an animal. The nucleic acids, and the proteins they
CC   encode may be used in the prevention, treatment and diagnosis of diseases
CC   associated with flea infestations. For example, the nucleic acids may be
CC   used to produce an HMT or HNC protein according to standard recombinant
CC   DNA methodology by inserting the nucleic acids into a host cell and
CC   culturing the cell to express the protein. The HMT and HNC nucleic acids
CC   may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect
CC   and quantitate the presence of cat flea or other homologous nucleic acid
CC   sequences in samples. They may also be used to study the expression and
CC   function of the proteins and their role in metabolism. The HMT and HNC
CC   proteins may be used as antigens in the production of specific
CC   antibodies, and in assays to identify modulators (agonists and
CC   antagonists) of HMT and/or HNC protein expression and activity. The
CC   anti-HMT/HNC protein antibodies and antagonists may also be used to
CC   downregulate protein expression and activity. The antibodies may also be
CC   used as diagnostic agents for detecting the presence of flea polypeptides
CC   in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The
CC   present sequence represents a cat flea HMT protein of the invention.
XX
SQ   Sequence   608 AA;

  Query Match             21.5%;  Score 595;  DB 21;  Length 608;
  Best Local Similarity   31.3%;  Pred. No. 6.6e-51;
  Matches  173;  Conservative  114;  Mismatches  190;  Indels   76;  Gaps   22;

Qy       6 SSLLQNSDLFTSDHASV-----VSMNLFVAL--LCACIVLGHLLEEN--RWVNESITALI 56
Db      59 NATIENPDPVLPEKGSAEQEQHSSMSIFFVLCVLALGILLIHFMLQTGFQYLPESIVVVF 118

Qy      57 IGLCTGVVILLLSGGKSSHLL---VFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIM 113
Db     119 LGALIGLIINLMSSKNIANWKNEEAFSPTAFFLVLLPPIIFESGYNLHKGNFFQNIGSIL 178

Qy     114 LFGAIGTLISCSIISFGAVKIFKHLDIDF-LDFGDYLAIGAIFAATDSVCTLQVL-SQDE 171
Db     179 VFAIFGTAISAFVVGAG-VYLLGMADVAYNLSFVESFAFGSLISAVDPVATVAIFHALDV 237

Qy     172 TPLLYSLVFGEGVVNDATSVVLFNAIQSFD---MTSFDPKI-GLHFIGNFLYLFLSSTFL 227
Db     238 DPVLNMLVFGESILNDAISIVLTTAVLESNNPLMTTAEAVVSGLN---RFCLMFFASAGI 294

Qy     228 GVGIGLLCAYIIKKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSH 287
Db     295 GVVFALISALLLKHVDLRKYPS-LELGMMLVFTYAPYVLAEGIHLSGIMAILFCGIVMSH 353

Qy     288 YTWHNVTESSRVTTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSIL 347
Db     354 YTHFNLSTVTQITMQQTMRTLAFIAETCVFAYLGMAIFSF------RHRVEPALVIWSIV 407

Qy     348 VGLILVGRAAFVFPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTS 407
Db     408 --LCLIGRAANIFPLSWLVNQFREH---KITKKMAFIMWFSGL-RGAISYAL-------- 453

Qy     408 GHTSLH------ENAIMITSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTS 461
Db     454 ---SLHLEFSDETRHVIITTTLIIVLCTTLIFGGATMPLLKFL-----QANKKTRSATRR 505

Qy     462 EPSSPKHFTVPLL---DNQPDSE--SDMITGPEVARPTALRMLLRTPTHTVHRYWRKF-- 514
Db     506 TRRQQKAITLSKTREWGSAIDSELLSELTTEEE----------RDVTFTQVRRGLEFIR 554

Qy     515 -DDSFMRPVFGGR 526
Db     555 LDHKYLRPFFTRR 567


RESULT 13
ABB65651
ID   ABB65651 standard; Protein; 727 AA.
XX
AC   ABB65651;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 23745.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL09754.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 23745; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   727 AA;

  Query Match             20.8%;  Score 575.5;  DB 22;  Length 727;
  Best Local Similarity   28.8%;  Pred. No. 7.9e-49;
  Matches  178;  Conservative   99;  Mismatches  196;  Indels  145;  Gaps   20;

Qy       5 LSSLLQNSDLFTSDHASVVSMNLFVALLCACIVLGHLLEENR--WVNESITALIIGLCTG 62
Db      71 LNHRIQSLDLL-----------VFVFLLALTVLTIWLFKHHRVSWLHETGLAVIYGLIVG 119

Qy      63 VVILLLSGGKSSHLL---------------------------------------------- 77
Db     120 AIIRY--AGTSATLVHMQVEPQGVPTYSDKLPPDTLWFRYPVNQTNGTKPPEGIKTYAYV 177

Qy      78 ----------------VFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGT 120
Db     178 FRGQVHDVDENEIDLKATFDPEVFFNIILPPIIFYAGYSLKKKYFFRNLGAILTFAIVGT 237

Qy     121 LISCSIIS---FGAVKIF-KHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPL-L 175
Db     238 TLSAFLIGGFMYGCVKLMPKYLSSSF-TFLDSLYFGALISPTDPLTILAIFNDLRVDVNL 296

Qy     176 YSLVFGEGVVNDATSVVLFNAIQS----------FDMTSFDPKIGLHFIGNFLYLFLSST 225
Db     297 YALVLGESVLNDAVAIVLSGAIQNYGEHYSNTGEFETTAF-----LRSLSDFFSIFLLSL 351

Qy     226 FLGVGIGLLCAYIIKKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVM 285
Db     352 MIGAAMGCLTA-LMTKFTRVRDFPLLESALFVLMSYSTFLLAEATELTGVVAVLFCGICQ 410
```

```
Qy     286 SHYTWHNVTESSRVTTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSS 345
           :|||::|::| ||  |:  |  |:|:|| ||| |:|     :  : | |:       :::
Db     411 AHYTYNNLSEDSRQRTKQIFELLNFLAENFIFSYIG-----VSMFTFPKHHFDAGFIITA 465

Qy     346 ILVGLILVGRAAFVFPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFT 405
             :   |  |||  |:|||:| |: :|    |||   | ::::||| |||:| |||
Db     466 FICAAI--GRAVNVYPLSWLLNIKRK---PKISTNFQHMLFFAGL-RGAMSFALAIR--- 516

Qy     406 TSGHTSLHENAIMITSTVTVVLFSTVVFGLMTKPLINLLLPP-----------HKQMPSG 454
             :|   |   |:|:|  :|:|: |: |     |:| |  |           : |: |
Db     517 ---NTVSDERQTMLTATSLIVIFTVVIQGGAANFLLNWLKIPVGVDDETEQLNNYQVHSS 573

Qy     455 HSSMTTSEPSSPKHFTVPL-------LDNQPDSESDMITGPEVARPTALRMLLRTPTH-- 505
             :   |        : |       ||   |  :  : |    |        |  :|
Db     574 DGYLQDVEGGGVNRNKMRLSGGTDTNLDTPVDGTNGSLGGASGGR--------RRNSHEK 625

Qy     506 -TVHRYWRKFDDSFMRPV 522
            : | |  ||  :|:|:
Db     626 AILARIWGNFDTKYMKPL 643
```

RESULT 14
AAY44017
ID AAY44017 standard; Protein; 633 AA.
XX
AC AAY44017;
XX
DT 18-JAN-2000 (first entry)
XX
DE Saccharomyces cerevisiae Na/H antiport transporter.
XX
KW Sodium; proton; antiport; transporter; salt tolerance; salt management;
KW transgenic plant; survival; soil; farming; accumulation; irrigation;
KW crop.
XX
OS Saccharomyces cerevisiae.
XX
PN WO9947679-A2.
XX
PD 23-SEP-1999.
XX
PF 18-MAR-1999; 99WO-CA00219.
XX
PR 18-MAR-1998; 98US-0078474.
PR 15-JAN-1999; 99US-0116111.
XX
PA (BLUM/) BLUMWALD E.
PA (APSE/) APSE M.
PA (SNED/) SNEDDEN W.
PA (AHAR/) AHARON G.
XX
PI Blumwald E, Apse M, Snedden W, Aharon G;
XX
DR WPI; 1999-571840/48.
XX
PT Nucleic acid molecules encoding sodium/proton transport polypeptides,
PT useful in genetic engineering salt tolerance in crop plants -
XX
PS Disclosure; Fig 8B; 93pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule encoding
CC a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment
CC and capable of increasing salt tolerance in a cell. This sequence
CC corresponds to an Na/H antiport transporter from the budding yeast
CC Saccharomyces cerevisiae. The Na/H transporter polypeptides provide
CC a means of intracellular salt management, particularly in plants. The
CC sequences are useful for producing transgenic plants that are capable
CC of surviving in soil with high salt levels that would normally inhibit
CC growth of the crop species. This would be useful in farming land in areas
CC that are generally considered unproductive through salt accumulation and
CC poor irrigation, e.g. in India, Australia, and prairies in USA or Canada.
CC Commercial crops, such as potatoes, tomatoes, brassica, cotton,
CC sunflower, strawberries, spinach, lettuce, rice, soybean, corn, wheat,
CC rye, barley, atriplex, sorghum, alfalfa, salicornia and others would
CC benefit from increased salt tolerance.
XX
SQ Sequence 633 AA;

```
  Query Match             20.5%;  Score 568;  DB 20;  Length 633;
  Best Local Similarity   27.2%;  Pred. No. 3.7e-48;
  Matches  166;  Conservative  117;  Mismatches  203;  Indels  124;  Gaps   19;

Qy       8 LLQNSDLFTSDHASV----VSMN----------LFVALLCACIVLG---HLLEEN-RWV 48
           || : ||  ||        | :|          ||: ||     |    :| ::  | |
Db      30 LLPSPDLPGSDDPIAGDPDVDLNPVTEEMFSSWALFIMLLLLISALWSSYYLTQKRIRAV 89

Qy      49 NESITALIIGLCTGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVN 108
           :|:: ::  |:  |::| :  |      : |:   ||  |||||| |:|::: :  || |
Db      90 HETVLSIFYGMVIGLIIRMSPGHYIQDTVTFNSSYFFNVLLPPIILNSGYELNQVNFFNN 149

Qy     109 FMTIMLFGAIGTLISCSIISFGAVKIFKHLDIDFLD--FGDYLAIGAIFAATDSVCTLQV 166
            ::|::|   || ||  :|   : |:  | :: :|  | | :::||  :||| |  | :
Db     150 MLSILIFAIPGTFISAVVIGI-ILYIWTFLGLESIDISFADAMSVGATLSATDPVTILSI 208

Qy     167 LSQDET-PLLYSLVFGEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSST 225
           :  :  | ||:::||| ::||| |:|:|   | |               | ||  |  |
Db     209 FNAYKVDPKLYTIIFGESLLNDAISIVMFETCQKFHGQPATFSSVFEGAGLFLMTFSVSL 268

Qy     226 FLGVGIGLLCAYIIKKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVM 285
           :|| ||:| | ::|  :  |:    |  |::|::| ||  :   ::|||::: |||| :
Db     269 LIGVLIGILVALLLKHTHIRRY-PQIESCLILLIAYESYFFSNGCHMSGIVSLLFCGITL 327

Qy     286 SHYTWHNVTESSRVTTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSS 345
           || ::|::  |::| :: |  |: ::| |||:|:|::     :  :     : | : |::
Db     328 KHYAYYNMSRRSQITIKYIFQLLARLSENFIFIYLGLELFTEVELVY----KPLLIIVAA 383

Qy     346 ILVGLILVGRAAFVFPLS------------FLSNLAKKNSS--DKISFRQQIIIWWAGL 390
           |    | | |   ||||              :| :  :| |  |:| :  |:: :||||
Db     384 I---SICVARWCAVFPLSQFVNWIYRVKTIRSMSGITGENISVPDEIPYNYQMMTFWAGL 440

Qy     391 MRGAVSIALAYN-----KFTTSGHTSLHENAIMITSTVTVVLFSTVVFGLMTKPLINLL- 444
           |||| :|||       |||             :: : : ||: : ::||  |  :: :|
Db     441 -RGAVGVALALGIQGEYKFT------------LLATVLVVVVLTVIIFGGTTAGMLEVLN 487

Qy     445 ------------------LPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMI 485
                            |    : :| |  |   | |              :| ||   |
Db     488 IKTGCISEEDTSDDEFDIEAPRAINLLNGSSIQTDLGPYSD--------NNSPDISIDQF 539

Qy     486 ---------------------------TGPEVARPTALRMLLRTPTHTV----HRYWRK 513
                                       | |  || :  :  ||    |:    ::::
Db     540 AVSSNKNLPNNISTTGGNTFGGLNETENTSPNPARSSMDKRNLRDKLGTIFNSDSQWFQN 599

Qy     514 FDDSFMRPVF 523
           ||:  ::|||
Db     600 FDEQVLKPVF 609
```

RESULT 15
AAY44016
ID AAY44016 standard; Protein; 569 AA.
XX
AC AAY44016;
XX
DT 18-JAN-2000 (first entry)
XX
DE Schizosaccharomyces pombe Na/H antiport transporter.
XX
KW Sodium; proton; antiport; transporter; salt tolerance; salt management;
KW transgenic plant; survival; soil; farming; accumulation; irrigation;
KW crop.
XX
OS Schizosaccharomyces pombe.
XX
PN WO9947679-A2.
XX
PD 23-SEP-1999.

```
XX
PF   18-MAR-1999;   99WO-CA00219.
XX
PR   18-MAR-1998;   98US-0078474.
PR   15-JAN-1999;   99US-0116111.
XX
PA   (BLUM/) BLUMWALD E.
PA   (APSE/) APSE M.
PA   (SNED/) SNEDDEN W.
PA   (AHAR/) AHARON G.
XX
PI   Blumwald E,  Apse M,  Snedden W,  Aharon G;
XX
DR   WPI; 1999-571840/48.
XX
PT   Nucleic acid molecules encoding sodium/proton transport polypeptides,
PT   useful in genetic engineering salt tolerance in crop plants -
XX
PS   Disclosure; Fig 8A; 93pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule encoding
CC   a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment
CC   and capable of increasing salt tolerance in a cell.  This sequence
CC   corresponds to an Na/H antiport transporter from the fission yeast
CC   Schizosaccharomyces pombe.  The Na/H transporter polypeptides provide
CC   a means of intracellular salt management, particularly in plants. The
CC   sequences are useful for producing transgenic plants that are capable
CC   of surviving in soil with high salt levels that would normally inhibit
CC   growth of the crop species. This would be useful in farming land in areas
CC   that are generally considered unproductive through salt accumulation and
CC   poor irrigation, e.g. in India, Australia, and prairies in USA or Canada.
CC   Commercial crops, such as potatoes, tomatoes, brassica, cotton,
CC   sunflower, strawberries, spinach, lettuce, rice, soybean, corn, wheat,
CC   rye, barley, atriplex, sorghum, alfalfa, salicornia and others would
CC   benefit from increased salt tolerance.
XX
SQ   Sequence   569 AA;

  Query Match             20.1%;  Score 556;  DB 20;  Length 569;
  Best Local Similarity   28.9%;  Pred. No. 5.1e-47;
  Matches  162;  Conservative  119;  Mismatches  181;  Indels   98;  Gaps   23;

Qy      24 SMNLFVALLCACIVLGHLL-------EENRWVNESITALIIGLCTGVVILLLSGGKSSHL 76
           |  ||: |:   :::| ||       :: | ::|:: :: :|:  |::| :  |    ::
Db      34 SWALFILLV---LLIGALLTSYYVQSKKIRAIHETVISVFVGMVVGLIIRVSPGLIIQNM 90

Qy      77 LVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIISFGA-VKIF 135
           : |    ||  |||||| |:|::: :  || |  ||: |   || ||   :: |  | ||
Db      91 VSFHSTYFFNVLLPPIILNSGYELHQSNFFRNIGTILTFAFAGTFISA--VTLGVLVYIF 148

Qy     136 KHLDIDFLD--FGDYLAIGAIFAATDSVCTLQVL-SQDETPLLYSLVFGEGVVNDATSVV 192
             |: : |   | : |::||  :||| |  | :  |      ||:::||| ::||| ::|
Db     149 SFLNFENLSMTFVEALSMGATLSATDPVTVLAIFNSYKVDQKLYTIIFGESILNDAVAIV 208

Qy     193 LFNAIQSFDMTSFDPKIGLHF------IGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGR 246
           :|  :| |   :      |||      || |:  |  |  :|| |||: | ::|  |  |
Db     209 MFETLQQFQGKT------LHFFTLFSGIGIFIITFFISLLIGVSIGLITALLLKYSYLRR 262

Qy     247 HSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFA 306
           : :  |  :::||:| ||  :  ::||:::: |||| : || : |::  ::::|:: |
Db     263 YPS-IESCIILLMAYTSYFFSNGCHMSGVVSLLFCGITLKHYAFFNMSYKAKLSTKYVFR 321

Qy     307 TLSFVAETFIFLYVGMDALD----IEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPL 362
           |: ::| |||:|:||        : |  |:  :   :|||::         |   ||||
Db     322 VLAQLSENFIFIYLGMSLFTQVDLVYKPIFILIT---TVAVTA--------SRYMNVFPL 370

Qy     363 SFLSNLAKK----NSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENA-I 417
           | | |   :    |  | | :  |::::|||| |||| :|||      :|     |||
Db     371 SNLLNKFHRQRNGNLIDHIPYSYQMMLFWAGL-RGAVGVALA------AGFEG--ENAQT 421

Qy     418 MITSTVTVVLFSTVVFGLMTKPLINLL----------------LPPHKQMPS---GHSS 457
           :  :|: ||: : ::||  |  :: :|                : | :| |     :|:
Db     422 LRATTLVVVVLTLIIFGGTTARMLEILHIETGVAADVDSDTEIGMLPWQQSPEFDLENSA 481

Qy     458 MTTSEPS----------SPKHF----TVPLLDNQPDSESDMITGPEVARPT-ALRMLLRT 502
           |  |: |          : :||      | |  :  |     | :  |  |     :
Db     482 MELSDASAEPVVVDQQFTTEHFDEGNIAPTLSKKVSS-----TFEQYQRAAGAFNQFFHS 536

Qy     503 PTHTVHRYWRKFDDSFMRPV 522
                 ::  :||:  ::||
Db     537 SRDDQAQWLTRFDEEVIKPV 556


Search completed: October 18, 2002, 12:25:19
Job time : 38 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: October 18, 2002, 12:24:10 ; Search time 21 Seconds
(without alignments)
2480.020 Million cell updates/sec

Title: US-09-830-123-2
Perfect score: 2768
Sequence: 1 MAFGLSSLLQNSDLFTSDHA.....FGGRGFVPFVAGSPVEQSPR 542

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1732.5	62.6	457	2	T01804	Na+/H+-exchanging
2	591.5	21.4	703	2	T26529	hypothetical prote
3	568	20.5	633	2	S69734	hypothetical prote
4	556	20.1	569	2	T37706	probable sodium/hy
5	552.5	20.0	629	2	T22848	hypothetical prote
6	485.5	17.5	813	2	A46748	Na+/H+-exchanging
7	482.5	17.4	809	2	A46747	Na+/H+-exchanging
8	471.5	17.0	832	2	A40205	Na+/H+-exchanging
9	468.5	16.9	831	2	B40204	Na+/H+-exchanging
10	468	16.9	759	2	A46188	cAMP-activated Na+
11	467.5	16.9	822	2	S30198	Na+/H+-exchanging
12	466	16.8	698	2	A57644	Na+/H+-exchanging
13	460.5	16.6	818	2	A48858	Na+/H+-exchanging
14	460	16.6	820	2	A40204	Na+/H+-exchanging
15	459.5	16.6	816	2	S16328	Na+/H+-exchanging
16	456.5	16.5	815	2	I57487	Na+/H+-exchanging
17	437.5	15.8	717	2	C40204	Na+/H+-exchanging
18	425.5	15.4	634	2	T33528	hypothetical prote
19	419	15.1	478	2	T18746	Na+/H+-exchanging
20	404	14.6	602	2	H88215	protein B0495.4 [i
21	397.5	14.4	798	2	T23539	hypothetical prote
22	389.5	14.1	651	2	T31869	hypothetical prote
23	376	13.6	375	2	B40205	Na+/H+-exchanging
24	375	13.5	660	2	T28016	hypothetical prote
25	335.5	12.1	609	2	S30910	Na+/H+-exchanging
26	319	11.5	520	2	AD2383	Na+/H+-exchanging
27	311.5	11.3	684	2	T16072	hypothetical prote
28	304	11.0	527	1	S75063	Na+/H+-exchanging
29	302	10.9	494	1	F69355	Na+/H+ antiporter
30	280.5	10.1	531	2	D96827	protein F20B17.4 [
31	275.5	10.0	528	2	T34941	probable Na+/H+ an
32	268	9.7	575	2	D96585	hypothetical prote
33	262.5	9.5	550	2	B98350	hypothetical prote
34	257.5	9.3	527	2	AE2932	Na+/H+ antiporter
35	245.5	8.9	682	2	AC1175	probable Na+/H+ an
36	243.5	8.8	682	2	AC1532	probable Na+/H+ an
37	243	8.8	458	2	H75278	probable Na+/H+ an
38	243	8.8	517	2	AC2070	Na+/H+ antiporter
39	235	8.5	490	2	F84154	Na+/H+ antiporter
40	234	8.5	565	2	H82610	sodium/proton exch
41	234	8.5	581	2	G83189	probable sodium/hy
42	226	8.2	650	2	AB1738	probable Na+/H+ an
43	226	8.2	650	2	AI1368	probable Na+/H+ an
44	226	8.2	680	2	F89832	conserved hypothet
45	224.5	8.1	540	1	S76231	Na+/H+-exchanging

ALIGNMENTS

RESULT 1
T01804
Na+/H+-exchanging protein 3 homolog A_TM021B04.4 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 22-Oct-1999
C;Accession: T01804
R;Dante, M.; Wamsley, P.; Gibson, A.
submitted to the EMBL Data Library, June 1997
A;Description: The sequence of A. thaliana TM021B04.
A;Reference number: Z14440
A;Accession: T01804
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-457 <DAN>
A;Cross-references: EMBL:AF007271; NID:g2191181; PID:g2191184; GSPDB:GN00063; ATSP:A_
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: ATSP:A_TM021B04.4
A;Map position: 5
A;Introns: 55/3; 95/2; 128/1; 148/1; 163/3; 240/3; 256/2; 286/3; 320/2; 400/3; 421/3

Query Match 62.6%; Score 1732.5; DB 2; Length 457;
Best Local Similarity 75.4%; Pred. No. 2.4e-123;
Matches 341; Conservative 44; Mismatches 46; Indels 21; Gaps 3;

```
Qy       5 LSSLLQN-SDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLCTGV 63
           | ||:     | ||||||||::|||||||||||||||||||||||:|||||||:||| |||
Db       2 LDSLVSKLPSLSTSDHASVVALNLFVALLCACIVLGHLLEENRWMNESITALLIGLGTGV 61

Qy      64 VILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLIS 123
            |||:| ||||||||||||||||||||||||||||||||||| ||:|||||||:||:||
Db      62 TILLISKGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFRNFVTIMLFGAVGTIIS 121

Qy     124 CSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVFGEG 183
           |:||| |  : || |||   | ||||||||||||||||||||||:|||||||||||||||
Db     122 CTIISLGVTQFFKKLDIGTFDLGDYLAIGAIFAATDSVCTLQVLNQDETPLLYSLVFGEG 181

Qy     184 VVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGI------------ 231
           |||||||||:|||||||:|  : :   | :|||||||| || ||  :
Db     182 VVNDATSVVVFNAIQSFDLTHLNHEAAFHLLGNFLYLFLLSTLLGAAVSPLFSSLPFFLT 241

Qy     232 GLLCAYIIKKLYFG--------RHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGI 283
           ||: ||:|||||||        |||||||||||||:||||::|||| ||||||||||||
Db     242 GLISAYVIKKLYFGRWPHINCHRHSTDREVALMMLMAYLSYMLAELFDLSGILTVFFCGI 301

Qy     284 VMSHYTWHNVTESSRVTTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAV 343
           |||||||||||||:||:|:|||||:|||||||||||||||||:|| | :: | |:||
Db     302 VMSHYTWHNVTESSRITTKHTFATLSFLAETFIFLYVGMDALDIDKWRSVSDTPGTSIAV 361

Qy     344 SSILVGLILVGRAAFVFPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNK 403
           ||||:||::|||||||||||||||||| |:||:|  |::|||:||||||||:||||||
Db     362 SSILMGLVMVGRAAFVFPLSFLSNLAKKNQSEKINFNMQVVIWWSGLMRGAVSMALAYNK 421
```

```
Qy    404 FTTSGHTSLHENAIMITSTVTVVLFSTVVFGL 435
          || :||| :  |||||||||:|| ||||||  |
Db    422 FTRAGHTDVRGNAIMITSTITVCLFSTVVRNL 453
```

RESULT 2
T26529
hypothetical protein Y18D10A.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T26529
R;Harris, B.
submitted to the EMBL Data Library, December 1998
A;Reference number: Z20226
A;Accession: T26529
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-703 <WIL>
A;Cross-references: EMBL:AL034393; PIDN:CAA22320.1; CESP:Y18D10A.6
A;Experimental source: clone Y18D10A
C;Genetics:
A;Gene: CESP:Y18D10A.6
A;Introns: 23/3; 56/1; 103/2; 246/3; 352/3; 492/2; 578/3

  Query Match 21.4%; Score 591.5; DB 2; Length 703;
  Best Local Similarity 34.8%; Pred. No. 6.8e-37;
  Matches 162; Conservative 78; Mismatches 181; Indels 45; Gaps 13;

```
Qy     27 LFVALLCACIVLGHLLEENRWVNESITALIIGLCTGVVILLLSGGKSSHLLVFSEDLFFI 86
          ||| :|   :|   :: :  |: ||:  : :|   |  ||  |    | :  | |:||:
Db    138 LFVIMLATLVVHMLIVSKIHWMPESLAIVALGALIG-SILSYSRRDWSEIEALSPDVFFL 196

Qy     87 YLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIISFGAVKIFKHLDIDF-LDF 145
          |||||||   : : |  || ||: |: |   || ||  :|  | : |   : : |   |
Db    197 VLLPPIIFENAYNLNKGYFFSNFVPILTFAIFGTTISAMVIGAG-LYILGAIGLIFEFTF 255

Qy    146 GDYLAIGAIFAATDSVCTLQVLSQDET-PLLYSLVFGEGVVNDATSVVLF-----NAIQS 199
          :  |  |: :| | | || :    :   ||| ||||| ::||| |:||      :|  |
Db    256 FECFAFAAMISAVDPVGTLAIFQAVKVESLLYMLVFGESMLNDAVSIVLAATALRHAKPS 315

Qy    200 FDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDREVALMMLM 259
          |:      |  |:  |  :|  |  |||||||  | : | :   | :   | ||:::
Db    316 FNSLPASEIITSAFV-TFTEMFFFSACLGVGIGLLSALLFKHVDL-RKTPSLEFALLLIF 373

Qy    260 SYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETFIFLY 319
          ||: |  ||   ||||: : |||| || :| |||:  :::| ||:| |:|||||  | |
Db    374 SYIPYGFAEALDLSGIMAILFCGISMSQFTRHNVSPIAQITFRHTFRTISFVAETSTFAY 433

Qy    320 VGMDALDIEKWKFVKNSQGLSVAVSSIL--VGLILVGRAAFVFPLSFLSNLAKKNSSDKI 377
          :||    |:          |: |   |   | | |:|||  ||||::| |  :|:   :|
Db    434 IGMAFFTIK----------LNFAPWLIFWSVVLCLLGRACNVFPLAYLVNQCRKDV--QI 481

Qy    378 SFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTVVFGLMT 437
          | : |||:|::| |||||  ||          : :|::|: : ::||:|:  |
Db    482 SMKNQIIMWFSG-MRGAVCFALVLYM-----DLDKEKKSILLTTVLFLILFTTIFLGGSA 535

Qy    438 KPLINLL--------------LPPHKQMPSGHSSMTTSEPSSPKHF 469
          | |: :                | :|:     |::  |:      |
Db    536 LPFISFINRCYPNERQRKRRRTPRNKESTGNSSALMMSKTQEMSFF 581
```

RESULT 3
S69734
hypothetical protein YDR456w - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 23-Mar-2001
C;Accession: S69734
R;Dietrich, F.S.
submitted to the EMBL Data Library, August 1995
A;Description: The sequence of S. cerevisiae lambda 3641 and cosmids 9461, 9831, and 941
A;Reference number: S69555
A;Accession: S69734
A;Molecule type: DNA
A;Residues: 1-633 <DIE>
A;Cross-references: EMBL:U33007; NID:g927685; PID:g927695; GSPDB:GN00004; MIPS:YDR456
C;Genetics:
A;Gene: MIPS:YDR456w
A;Map position: 4R
C;Superfamily: hypothetical protein yvgP

  Query Match 20.5%; Score 568; DB 2; Length 633;
  Best Local Similarity 27.2%; Pred. No. 3.6e-35;
  Matches 166; Conservative 117; Mismatches 203; Indels 124; Gaps 19;

```
Qy      8 LLQNSDLFTSDHASV----VSMN----------LFVALLCACIVLG---HLLEEN-RWV 48
          || : ||  ||        | :|          ||: ||     |    :| ::  | |
Db     30 LLPSPDLPGSDDPIAGDPDVDLNPVTEEMFSSWALFIMLLLLISALWSSYYLTQKRIRAV 89

Qy     49 NESITALIIGLCTGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVN 108
          :|:: ::  |:  |::| :  |      : |:   ||  |||||| |:|::: :  || |
Db     90 HETVLSIFYGMVIGLIIRMSPGHYIQDTVTFNSSYFFNVLLPPIILNSGYELNQVNFFNN 149

Qy    109 FMTIMLFGAIGTLISCSIISFGAVKIFKHLDIDFLD--FGDYLAIGAIFAATDSVCTLQV 166
          ::|::|   || ||  :|    : |:  | :: :|  | | :::||  :||| |  | :
Db    150 MLSILIFAIPGTFISAVVIGI-ILYIWTFLGLESIDISFADAMSVGATLSATDPVTILSI 208

Qy    167 LSQDET-PLLYSLVFGEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSST 225
          :  :  | ||:::||| ::||| |:|:|   | |               | ||  |  |
Db    209 FNAYKVDPKLYTIIFGESLLNDAISIVMFETCQKFHGQPATFSSVFEGAGLFLMTFSVSL 268

Qy    226 FLGVGIGLLCAYIIKKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVM 285
          :|| ||:| | ::|  :  |:    |  |::|::| ||  :   ::|||::: |||| :
Db    269 LIGVLIGILVALLLKHTHIRRY-PQIESCLILLIAYESYFFSNGCHMSGIVSLLFCGITL 327

Qy    286 SHYTWHNVTESSRVTTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSS 345
          || ::|::  |::| :: |  |: ::| |||:|:|::     :  :    : | : |::
Db    328 KHYAYYNMSRRSQITIKYIFQLLARLSENFIFIYLGLELFTEVELVY----KPLLIIVAA 383

Qy    346 ILVGLILVGRAAFVFPLS------------FLSNLAKKNSS--DKISFRQQIIIWWAGL 390
          |    | | |   |||||              :| :  :| |  |:| :  |:: :||||
Db    384 I---SICVARWCAVFPLSQFVNWIYRVKTIRSMSGITGENISVPDEIPYNYQMMTFWAGL 440

Qy    391 MRGAVSIALAYN-----KFTTSGHTSLHENAIMITSTVTVVLFSTVVFGLMTKPLINLL- 444
          |||| :|||       |||             :: : : ||: : ::||  |  :: :|
Db    441 -RGAVGVALALGIQGEYKFT------------LLATVLVVVVLTVIIFGGTTAGMLEVLN 487

Qy    445 ------------------LPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMI 485
                             |    : :| |  |   | |          :| ||   |
Db    488 IKTGCISEEDTSDDEFDIEAPRAINLLNGSSIQTDLGPYSD--------NNSPDISIDQF 539

Qy    486 ---------------------------TGPEVARPTALRMLLRTPTHTV----HRYWRK 513
                                      | |  || :  :  ||    |:     ::::
Db    540 AVSSNKNLPNNISTTGGNTFGGLNETENTSPNPARSSMDKRNLRDKLGTIFNSDSQWFQN 599

Qy    514 FDDSFMRPVF 523
          ||:  ::|||
Db    600 FDEQVLKPVF 609
```

RESULT 4
T37706
probable sodium/hydrogen exchanger - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C;Accession: T37706
R;Murphy, L.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1997
A;Reference number: Z21738
A;Accession: T37706
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-569 <MUR>
A;Cross-references: EMBL:Z97208; PIDN:CAB10103.1; GSPDB:GN00066; SPDB:SPAC15A10.06
A;Experimental source: strain 972h-; cosmid c15A10
C;Genetics:

```
A;Gene: SPDB:SPAC15A10.06
A;Map position: 1
A;Introns: 11/1; 116/3; 356/3
C;Superfamily: hypothetical protein yvgP

  Query Match           20.1%;  Score 556;  DB 2;  Length 569;
  Best Local Similarity  28.9%;  Pred. No. 2.5e-34;
  Matches  162;  Conservative  119;  Mismatches  181;  Indels   98;  Gaps   23;

Qy      24 SMNLFVALLCACIVLGHLL-------EENRWVNESITALIIGLCTGVVILLLSGGKSSHL 76
           |  ||: |:   :::| ||       :: | ::|:: :: :|:  |::| :  |    ::
Db      34 SWALFILLV---LLIGALLTSYYVQSKKIRAIHETVISVFVGMVVGLIIRVSPGLIIQNM 90

Qy      77 LVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIISFGA-VKIF 135
           : |    ||  |||||| |:|::: :  || |  ||: |   || ||   :: |  | ||
Db      91 VSFHSTYFFNVLLPPIILNSGYELHQSNFFRNIGTILTFAFAGTFISA--VTLGVLVYIF 148

Qy     136 KHLDIDFLD--FGDYLAIGAIFAATDSVCTLQVL-SQDETPLLYSLVFGEGVVNDATSVV 192
             |: : |   | : |::||  :||| |  | : |      ||:::||| ::||| ::|
Db     149 SFLNFENLSMTFVEALSMGATLSATDPVTVLAIFNSYKVDQKLYTIIFGESILNDAVAIV 208

Qy     193 LFNAIQSFDMTSFDPKIGLHF------IGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGR 246
           :|  :| |   :      |||      || |:  |  |  :|| ||| : ::|  |  |
Db     209 MFETLQQFQGKT------LHFFTLFSGIGIFIITFFISLLIGVSIGLITALLLKYSYLRR 262

Qy     247 HSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFA 306
           : :  |  :::||:| ||  :   ::||:::: |||| : || : |::  ::::|:: |
Db     263 YPS-IESCIILLMAYTSYFFSNGCHMSGVVSLLFCGITLKHYAFFNMSYKAKLSTKYVFR 321

Qy     307 TLSFVAETFIFLYVGMDALD----IEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPL 362
           |: ::| |||:|:|| |       : |  |:  :  :|||::          |  ||||
Db     322 VLAQLSENFIFIYLGMSLFTQVDLVYKPIFILIT---TVAVTA--------SRYMNVFPL 370

Qy     363 SFLSNLAKK----NSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENA-I 417
           | | |   :    |  | | :  |::::|||| |||| :|||      :|     |||
Db     371 SNLLNKFHRQRNGNLIDHIPYSYQMMLFWAGL-RGAVGVALA------AGFEG--ENAQT 421

Qy     418 MITSTVTVVLFSTVVFGLMTKPLINLL----------------LPPHKQMPS---GHSS 457
           :  :|: ||: : ::||  |  :: :|                 : | :| |     :|:
Db     422 LRATTLVVVVLTLIIFGGTTARMLEILHIETGVAADVDSDTEIGMLPWQQSPEFDLENSA 481

Qy     458 MTTSEPS----------SPKHF----TVPLLDNQPDSESDMITGPEVARPT-ALRMLLRT 502
           |  |: |          : :||      | |  :  |     |  : |  |       :
Db     482 MELSDASAEPVVVDQQFTTEHFDEGNIAPTLSKKVSS-----TFEQYQRAAGAFNQFFHS 536

Qy     503 PTHTVHRYWRKFDDSFMRPV 522
               ::  :||:  ::||
Db     537 SRDDQAQWLTRFDEEVIKPV 556


RESULT 5
T22848
hypothetical protein F57C7.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T22848
R;White, S.
submitted to the EMBL Data Library, February 1996
A;Reference number: Z19625
A;Accession: T22848
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-629 <WIL>
A;Cross-references: EMBL:Z69646; PIDN:CAA93476.1; GSPDB:GN00028; CESP:F57C7.2
A;Experimental source: clone F57C7
C;Genetics:
A;Gene: CESP:F57C7.2
A;Map position: X
A;Introns: 33/3; 81/1; 119/2; 155/3; 207/2; 239/1; 274/2; 304/1; 354/1; 386/3; 421/3; 45

  Query Match           20.0%;  Score 552.5;  DB 2;  Length 629;
  Best Local Similarity  35.1%;  Pred. No. 5.3e-34;
  Matches  158;  Conservative   77;  Mismatches  162;  Indels   53;  Gaps   16;

Qy      78 VFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIISFGA---VKI 134
           ||| ::||  |:|||||||: :||: || |  :|:    ||| |||    ||    : :
Db     157 VFSPEVFFNMLIPPIIFNAGYSLKKRHFFRNIGSILAIVFIGTTISC----FGTGCLMFV 212

Qy     135 FKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPL-LYSLVFGEGVVNDATSVVL 193
           |  :     | : |  ||: :||| |  : | :      |::|:|||  :||| ::||
Db     213 FTSIFQMGYSFKELLFFGALISATDPVTIISVFNDMNVEADLFALIFGESALNDAVAIVL 272

Qy     194 FNAIQSFDMTSFDPKIGLHFIGN----FLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHST 249
             |::|  :|    | |   |:    |  :|  |  ||  || : |:: |      |:
Db     273 SEVIENFSTSS--EAITLQDFGSAIAGFAGVFFGSLMLGFMIGCMNAFLTKMTLISEHAL 330

Qy     250 DREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLS 309
           | :| :|:||:|:::||:  |:||::| ||||  :|||::|::: |:  |:| |  :|
Db     331 -LESSLFVLISYISFLVAEVCGLTGIVSVLFCGIAQAHYTYNNLSDESQSNTKHFFHMVS 389

Qy     310 FVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILV--GRAAFVFPLSFLSN 367
           |: |:||| |:|:         || |:|  |   | :|  || :   || ||:|||:| |
Db     390 FIMESFIFCYIGVSV-------FVTNNQRWS--FSFLLFSLISITASRALFVYPLSWLLN 440

Qy     368 LAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVL 427
           : ::    ||  | | :| :||| |||:: |||       |  :  ||  || :| | |:
Db     441 IRRR---PKIPKRYQHMILFAGL-RGAMAFALA-------GRNTSTENRQMIFATTTAVV 489

Qy     428 FSTV-VFGLMTKPLINLLLPPH-KQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMI 485
           || | | :|  :|: |   | |      :  |  |||                |  ||:
Db     490 IVTVLVNGGLTSWMIDYLQIKHGKDAIEEGQRLENSMSSSP----------ADQHSDLD 538

Qy     486 TGPEVARPTALRMLLRTPTHTVHRYWRKFD 515
             |     |    :     : | |  ||
Db     539 ESVPVTMSPGLNPWDKA---FLPRKWYHFD 565


RESULT 6
A46748
Na+/H+-exchanging protein NHE-2 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 05-Nov-1999
C;Accession: A46748; A47449
R;Wang, Z.; Orlowski, J.; Shull, G.E.
J. Biol. Chem. 268, 11925-11928, 1993
A;Title: Primary structure and functional expression of a novel gastrointestinal isof
A;Reference number: A46748; MUID:93280160
A;Accession: A46748
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-813 <WAN>
A;Cross-references: GB:L11236; NID:g205318; PIDN:AAA72350.1; PID:g205319
R;Collins, J.F.; Honda, T.; Knobel, S.; Bulus, N.M.; Conary, J.; DuBois, R.; Ghishan,
Proc. Natl. Acad. Sci. U.S.A. 90, 3938-3942, 1993
A;Title: Molecular cloning, sequencing, tissue distribution, and functional expressio
A;Reference number: A47449; MUID:93248205
A;Accession: A47449
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 117-813 <COL>
A;Experimental source: intestine
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:130778, NCBIP:130779)

  Query Match           17.5%;  Score 485.5;  DB 2;  Length 813;
  Best Local Similarity  28.0%;  Pred. No. 8.3e-29;
  Matches  148;  Conservative   86;  Mismatches  215;  Indels   79;  Gaps   18;

Qy      14 LFTSDHASV---VSMNLFVALLCACIVLGHLLEE-NRWVNESITALIIGLCTGVVILLLS 69
           :|| |: |     : |:: |      :  ||  :    | ||   :::||  | :|
Db      71 VFTLDYPHVQIPFEITLWILLASLAKIGFHLYHKLPTIVPESCLLIMVGLLLGGIIF--- 127

Qy      70 GGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTL-----ISC 124
           |           |:||:|||||: :||: :  : || |  ||  :  :|||     |
Db     128 GVDEKSPPAMKTDVFFLYLLPPIVLDAGYFMPTRPFFENLGTIFWYAVVGTLWNSIGIGL 187
```

```
Qy    125 SIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQ-DETPLLYSLVFGEG 183
          |:     :: |   ||  |     |  |:: :| | |  | |         || |||||
Db    188 SLFGICQIEAFGLSDITLL---QNLLFGSLISAVDPVAVLAVFENIHVNEQLYILVFGES 244

Qy    184 VVNDATSVVLFNAIQSF----DMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYII 239
          ::||| :|||:|  :||     : : |   |   | ||  : :    :|: :| : |:
Db    245 LLNDAVTVVLYNLFKSFCQMKTIQTVDVFAG---IANFFVVGIGGVLIGILLGFIAAFTT 301

Qy    240 KKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRV 299
          :   |  :    |   : | ||||||  ||:|:||||: :  | : |: |   ||:: |
Db    302 R---FTHNIRVIEPLFVFLYSYLSYITAEMFHLSGIMAITACAMTMNKYVEENVSQKSYT 358

Qy    300 TTRHSFATLSFVAETFIFLYVGMDALDIE---KWKFVKNSQGLSVAVSSILVGLILVGRA 356
          | ::    || |:|| ||:::|:  :       | ||              :   |: ||
Db    359 TIKYFMKMLSSVSETLIFIFMGVSTVGKNHEWNWAFV----------CFTLAFCLIWRA 407

Qy    357 AFVFPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENA 416
            || |: : |  :   :  ::|: | || : || |||:  || :        |
Db    408 LGVFVLTQVINWFR---TIPLTFKDQFIIAYGGL-RGAICFALVF----LLPATVFPRKK 459

Qy    417 IMITSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDN 476
          : ||: : |: |:  : |:  :||:  |        |       ||    : |     |:
Db    460 LFITAAIVVIFFTVFILGITIRPLVEFL----DVKRSNKKQQAVSEEIHCRFF-----DH 510

Qy    477 QPDSESDMITGPEVARPTALRMLLRTPTHTVHRYWR----KFDDSFMR 520
                |:                   |  | :||    |||| ::|
Db    511 VKTGIEDVC-----------------GHWGHNFWRDKFKKFDDKYLR 540
```

RESULT 7
A46747
Na+/H+-exchanging protein NHE-2 - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 28-Feb-1997
C;Accession: A46747
R;Tse, C.M.; Levine, S.A.; Yun, C.H.; Montrose, M.H.; Little, P.J.; Pouyssegur, J.; Donc
J. Biol. Chem. 268, 11917-11924, 1993
A;Title: Cloning and expression of a rabbit cDNA encoding a serum-activated ethylisoprop
A;Reference number: A46747; MUID:93280159
A;Accession: A46747
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-809 <TSE>
A;Experimental source: ileal villus cells
A;Note: sequence extracted from NCBI backbone (NCBIN:133350, NCBIP:133351)

Query Match 17.4%; Score 482.5; DB 2; Length 809;
Best Local Similarity 27.8%; Pred. No. 1.4e-28;
Matches 147; Conservative 86; Mismatches 216; Indels 79; Gaps 18;

```
Qy     14 LFTSDHASV---VSMNLFVALLCACIVLGHLLEE-NRWVNESITALIIGLCTGVVILLLS 69
          :|| |:  |     : |:: |     :  ||  :    | ||   :::||  | :|
Db     70 VFTLDYPHVQIPFEITLWILLASLAKIGFHLYHKLPTIVPESCLLIMVGLLLGGIIF--- 126

Qy     70 GGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTL-----ISC 124
          |          |:||:|||||: :||: :  : || |  ||  :  :|||     |
Db    127 GVDEKSPPAMKTDVFFLYLLPPIVLDAGYFMPTRPFFENLGTIFWYAVVGTLWNSIGIGV 186

Qy    125 SIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQ-DETPLLYSLVFGEG 183
          |:     :: |   ||  |     |  |:: :| | |  | |         || |||||
Db    187 SLFGICQIEAFGLSDITLL---QNLLFGSLISAVDPVAVLAVFENIHVNEQLYILVFGES 243

Qy    184 VVNDATSVVLFNAIQSF----DMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYII 239
          ::||| :|||:|  :||     : : |   |   | ||  : :    :|: :| : |:
Db    244 LLNDAVTVVLYNLFKSFCQMKTIETIDVFAG---IANFFVVGIGGVLIGIFLGFIAAFTT 300

Qy    240 KKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRV 299
          :   |  :    |   : | ||||||  ||:|:||||: :  | : |: |   ||:: |
Db    301 R---FTHNIRVIEPLFVFLYSYLSYITAEMFHLSGIMAITACAMTMNKYVEENVSQKSYT 357

Qy    300 TTRHSFATLSFVAETFIFLYVGMDALDIE---KWKFVKNSQGLSVAVSSILVGLILVGRA 356
          | ::    || |:|| ||:::|:  :       | ||              :   |: ||
Db    358 TIKYFMKMLSSVSETLIFIFMGVSTVGKNHEWNWAFV----------CFTLAFCLIWRA 406

Qy    357 AFVFPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENA 416
            || |: : |  :   :  ::|: | || : || |||:  || :
Db    407 LGVFVLTRVINWFR---TIPLTFKDQFIIAYGGL-RGAICFALVF----LLPAAVFPRKK 458

Qy    417 IMITSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDN 476
          : ||: : |: |:  : |:  :||:  |        |       ||    : |     |:
Db    459 LFITAAIVVIFFTVFILGITIRPLVEFL----DVKRSNKKQQAVSEEIHCRFF-----DH 509

Qy    477 QPDSESDMITGPEVARPTALRMLLRTPTHTVHRYWR----KFDDSFMR 520
                |:                   |  | :||    |||| ::|
Db    510 VKTGIEDVC-----------------GHWGHNFWRDKFKKFDDKYLR 539
```

RESULT 8
A40205
Na+/H+-exchanging protein 3 - rabbit
N;Alternate names: Na+/H+ antiporter
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 05-Nov-1999
C;Accession: A40205
R;Tse, C.M.; Brant, S.R.; Walker, M.S.; Pouyssegur, J.; Donowitz, M.
J. Biol. Chem. 267, 9340-9346, 1992
A;Title: Cloning and sequencing of a rabbit cDNA encoding an intestinal and kidney-sp
A;Reference number: A40205; MUID:92250540
A;Accession: A40205
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-832 <TSE>
A;Cross-references: GB:M87007; NID:g165548; PIDN:AAA31420.1; PID:g165549
C;Keywords: transmembrane protein

Query Match 17.0%; Score 471.5; DB 2; Length 832;
Best Local Similarity 28.5%; Pred. No. 9.8e-28;
Matches 148; Conservative 92; Mismatches 197; Indels 83; Gaps 19;

```
Qy     22 VVSMNLFVALLCACIV-LGHLLEENRWVNESITALIIGLCTGVVILLLSGGKSSHLLVF- 79
          :::: : || |   :  | |  :    | ||   :::||  | ::|      : |:  |
Db     56 IIALWVLVASLAKIVFHLSH--KVTSVVPESALLIVLGLVLGGIVL-----AADHIASFT 108

Qy     80 -SEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIISFGAVKIF--- 135
           :  :|| ||||||: :||: :  : || |  :|:|:  :||: : :        :|
Db    109 LTPTVFFFYLLPPIVLDAGYFMPNRLFFSNLGSILLYAVVGTVWNAATTGLSLYGVFLSG 168

Qy    136 --KHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQ-DETPLLYSLVFGEGVVNDATSVV 192
             | |  |||    |  |:: || | |  | |  :      :|: :|||| ::||| :||
Db    169 IMGELKIGLLDF---LLFGSLIAAVDPVAVLAVFEEVHVNEVLFIIVFGESLLNDAVTVV 225

Qy    193 LFNAIQSFDMTSFDPKIGLHF---IGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHST 249
          |:|  |||      |    |:      | :|  : |  | :||     | : : :   | :|
Db    226 LYNVFQSFVTLGGDKVTGVDCVKGIVSFFVVSLGGTLVGVVFAFLLSLVTR---FTKHVR 282

Qy    250 DREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLS 309
             |    : ::|||||: :|:  || || : ||||     |    |::| |  | |::   |:
Db    283 VIEPGFVFIISYLSYLTSEMLSLSSILAITFCGICCQKYVKANISEQSATTVRYTMKMLA 342

Qy    310 FVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVG--RAAFVFPLSFLSN 367
             ||| ||:::|: |:|    | :           : :|: |: |    ||  |   ::| |
Db    343 SGAETIIFMFLGISAVDPLIWTW---------NTAFVLLTLLFVSVFRAIGVVLQTWLLN 393

Qy    368 LAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAY----NKFTTSGHTSLHENAIMITSTV 423
             :    : |    |::: : || ||||: ||      ||       : |  : :::|:
Db    394 RYRMVQLELI---DQVVMSYGGL-RGAVAFALVALLDGNK--------VKEKNLFVSTTI 441

Qy    424 TVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESD 483
             || |: :  ||  |||:  |    |   | |     :|    : |   |      | :
Db    442 IVVFFTVIFQGLTIKPLVQWL----KVKRSEHREPKLNEKLHGRAFDHIL------SAIE 491

Qy    484 MITGPEVARPTALRMLLRTPTHTVHRY----WRKFDDSFM 519
             |:|                 | |    |  ||  |:
Db    492 DISG----------------QIGHNYLRDKWANFDRRFL 514
```

```
RESULT 9
B40204
Na+/H+-exchanging protein 3 - rat
N;Alternate names: Na+/H+ antiporter
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 05-Nov-1999
C;Accession: B40204
R;Orlowski, J.; Kandasamy, R.A.; Shull, G.E.
J. Biol. Chem. 267, 9331-9339, 1992
A;Title: Molecular cloning of putative members of the Na/H exchanger gene family. cDNA c
rally related proteins.
A;Reference number: A40204; MUID:92250539
A;Accession: B40204
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-831 <ORL>
A;Cross-references: GB:M85300; NID:g205706; PIDN:AAA41702.1; PID:g205707
C;Keywords: transmembrane protein

  Query Match             16.9%;  Score 468.5;  DB 2;  Length 831;
  Best Local Similarity   27.7%;  Pred. No. 1.6e-27;
  Matches  146;  Conservative   89;  Mismatches  200;  Indels   93;  Gaps   18;

Qy      22 VVSMNLFVALLCACIV-LGHLLEENRWVNESITALIIGLCTGVVILLLSGGKSSHLLVF- 79
           :::: : || |   :  | |  :    | ||   :::||  | ::       : |:  |
Db      54 IIALWILVASLAKIVFHLSH--KVTSVVPESALLIVLGLVLGGIV-----WAADHIASFT 106

Qy      80 -SEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIISFGAVKIF--- 135
            :  ||| ||||||: :||: :  : || |  ||:|:  |||: : :         :|
Db     107 LTPTLFFFYLLPPIVLDAGYFMPNRLFFGNLGTILLYAVIGTIWNAATTGLSLYGVFLSG 166

Qy     136 --KHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQ-DETPLLYSLVFGEGVVNDATSVV 192
             | |  |||   |  |:: || | |  | |  :     :|: :|||| ::||| :||
Db     167 LMGELKIGLLDF---LLFGSLIAAVDPVAVLAVFEEVHVNEVLFIIVFGESLLNDAVTVV 223

Qy     193 LFNAIQSFDMTSFDPKIGLHF---IGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHST 249
           |:|  :||     |   |:     | :|  : |  | :||    | : : :   | :|
Db     224 LYNVFESFVTLGGDAVTGVDCVKGIVSFFVVSLGGTLVGVIFAFLLSLVTR---FTKHVR 280

Qy     250 DREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLS 309
             |   : ::|||||: :|:  || || : ||||    |   |::| |  | |::   |:
Db     281 IIEPGFVFVISYLSYLTSEMLSLSAILAITFCGICCQKYVKANISEQSATTVRYTMKMLA 340

Qy     310 FVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLA 369
             ||| ||:::|: |:|   | :       : |   : :  | | ||  |   ::: |
Db     341 SGAETIIFMFLGISAVDPVIWTW-------NTAFVLLTLVFISVYRAIGVVLQTWILNRY 393

Qy     370 KKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFS 429
           :   : |    |::: : || ||||: ||           : |  : :::|: || |:
Db     394 RMVQLETI---DQVVMSYGGL-RGAVAYALV----VLLDEKKVKEKNLFVSTTLIVVFFT 445

Qy     430 TVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPK-----------HFTVPLLDNQP 478
             :  ||  |||:  |              :  ||   ||           |    : |
Db     446 VIFQGLTIKPLVQWL------------KVKRSEQREPKLNEKLHGRAFDHILSAIED--- 490

Qy     479 DSESDMITGPEVARPTALRMLLRTPTHTVHRY----WRKFDDSFMRPV 522
             |:|                    | |    |  ||  |:  |
Db     491 ------ISG----------------QIGHNYLRDKWSNFDRKFLSKV 515


RESULT 10
A46188
cAMP-activated Na+/H+-exchanging protein betaNHE - rainbow trout
C;Species: Oncorhynchus mykiss (rainbow trout)
C;Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 28-Feb-1997
C;Accession: A46188
R;Borgese, F.; Sardet, C.; Cappadoro, M.; Pouyssegur, J.; Motais, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 6765-6769, 1992
A;Title: Cloning and expression of a cAMP-activated Na+/H+ exchanger: evidence that the
A;Reference number: A46188; MUID:92357712
A;Accession: A46188
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-759 <BOR>
A;Experimental source: cephalic kidney hematopoietic tissues
A;Note: sequence extracted from NCBI backbone (NCBIP:113525)

  Query Match             16.9%;  Score 468;  DB 2;  Length 759;
  Best Local Similarity   28.2%;  Pred. No. 1.6e-27;
  Matches  148;  Conservative  100;  Mismatches  184;  Indels   92;  Gaps   22;

Qy      31 LLCACIVLG-HLLEE-NRWVNESITALIIGLCTGVVILLLSGGKSSHLLVFSEDLFFIYL 88
           ||   : || ||:   :  | ||   :::||  | :| ::  |:   :|     |||: |
Db      78 LLALLMKLGFHLIPRLSAVVPESCLLIVVGLLVGGLIKVI--GEEPPVL--DSQLFFLCL 133

Qy      89 LPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLIS----------CSIISFGAVKIFKH 137
           ||||| :||: :  : |  |  ||::|  |||| :           | | | |
Db     134 LPPIILDAGYFLPIRPFTENVGTILVFAVIGTLWNAFFMGGLLYALCQIESVG------- 186

Qy     138 LDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDE-TPLLYSLVFGEGVVNDATSVVLFNA 196
             :  :|    |  |:| :| | |  | |  :     |:: ||||| ::||| :|||:|
Db     187 --LSGVDLLACLLFGSIVSAVDPVAVLAVFEEIHINELVHILVFGESLLNDAVTVVLYNL 244

Qy     197 IQSFD----MTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDRE 252
           : |      :|  |  :|   :  |  : |     :|   | | |:  :   |  |:   |
Db     245 FEEFSKVGTVTVLDVFLG---VVCFFVVSLGGVLVGAIYGFLAAFTSR---FTSHTRVIE 298

Qy     253 VALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVA 312
             : | ||::|: :|:|:|||||: :  ||:||  |   |::  |  | ::     | |:
Db     299 PLFVFLYSYMAYLSSEMFHLSGIMALIACGVVMRPYVEANISHKSYTTIKYFLKMWSSVS 358

Qy     313 ETFIFLYVGMDAL---DIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLA 369
           || ||:::|:  :       | ||          :::::  | || |   |  |:|: |
Db     359 ETLIFIFLGVSTVAGPHAWNWTFV---------ITTVI--LCLVSRVLGVIGLTFIIN-- 405

Qy     370 KKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFS 429
           |    |:: : | |: : || |||:: :| |    :: |   :  : :|: :||: |:
Db     406 -KFRIVKLTKKDQFIVAYGGL-RGAIAFSLGY--LLSNSHQMRN---LFLTAIITVIFFT 458

Qy     430 TVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMITGPE 489
           | |:  :||: ||     |:          |:||  :       ||:       ::|| |
Db     459 VFVQGMTIRPLVELLAVKKKK---------ESKPSINEEIHTEFLDH-------LLTGVE 502

Qy     490 VARPTALRMLLRTPTHTVHRYWR----KFDDSFM-RPVFGGRGF 528
                         |  | :|:    :|: ::: | :  |  |
Db     503 -----------GVCGHYGHYHWKEKLNRFNKTYVKRWLIAGENF 535


RESULT 11
S30198
Na+/H+-exchanging protein - Chinese hamster
N;Alternate names: Na+/H+ antiporter; Na+/H+ exchanger
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Nov-1999
C;Accession: S30198; S29541
R;Counillon, L.; Pouyssegur, J.
Biochim. Biophys. Acta 1172, 343-345, 1993
A;Title: Nucleotide sequence of the Chinese hamster Na(+)/H(+) exchanger NHE1.
A;Reference number: S30198; MUID:93192332
A;Accession: S30198
A;Molecule type: mRNA
A;Residues: 1-822 <COU>
A;Cross-references: EMBL:X68970; NID:g49472; PIDN:CAA48771.1; PID:g49473
C;Genetics:
A;Gene: NHE1
C;Keywords: glycoprotein; ion transport; membrane protein
F;76,374,414/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             16.9%;  Score 467.5;  DB 2;  Length 822;
  Best Local Similarity   28.2%;  Pred. No. 1.9e-27;
  Matches  144;  Conservative  107;  Mismatches  189;  Indels   71;  Gaps   21;

Qy      25 MNLFVALLCACIVLGHLLEE-NRWVNESITALIIGLCTGVVILLLSGGKSSHLLVFSEDL 83
           ::|:: | |   :  |::   :  | ||   :::||  |   |:   |    |::   |    |:
Db     109 ISLWILLACLMKIGFHVIPTISSIVPESCLLIVVGLLVGG--LIKGVGETPPFL--QSDV 164
```

```
Qy      84 FFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIIS--FGAVKIFKHLDID 141
           ||::||||||| :||: :  :||  |  ||:::|  :||| :   :     || :     |:
Db     165 FFLFLLPPIILDAGYFLPLRQFTENLGTILIFAVVGTLWNAFFLGGLLYAVCLVGGEQIN 224

Qy     142 FLDFGDYLAIGAIFAATDSVCTLQVLSQDE-TPLLYSLVFGEGVVNDATSVVLFNAIQSF 200
           :   | |  |:| :| | |  : |  :     ||: ||||| ::||| :|||::  : |
Db     225 NIGLLDTLLFGSIISAVDPVAVVAVFEEIHINELLHILVFGESLLNDAVTVVLYHLFEEF 284

Qy     201 DMTSFDPKIGLH--FIG--NFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDREVALM 256
              ::|  ||:   |:|  :|  : |   |:||  |:: |:  :   |  |    |   :
Db     285 --ANYD-SIGISDIFLGFLSFFVVALGGVFVGVVYGVIAAFTSR---FTSHIRVIEPLFV 338

Qy     257 MLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETFI 316
           | ||::|: ||||:|||| :   |:||  |   |::  |  | ::     | |:|| |
Db     339 FLYSYMAYLSAELFHLSGIMALIASGVVMRPYVEANISHKSHTTIKYFLKMWSSVSETLI 398

Qy     317 FLYVGMDALDIE---KWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKKNS 373
           |:::|:  :       | ||         :|::|    |: |   |  |::  |   |
Db     399 FIFLGVSTVAGSHQWNWTFV---------ISTLL--FCLIARVLGVLVLTWFIN---KFR 444

Qy     374 SDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTVVF 433
             |:: : | || : || |||:: :| |       |  : :  : :|: :||: |:  |
Db     445 IVKLTPKDQFIIAYGGL-RGAIAFSLGY--LMDKKHFPMCD--LFLTAIITVIFFTVFVQ 499

Qy     434 GLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMITGPEVARP 493
           |:  :||::||    ||          :: |  :      ||:       ::|| |
Db     500 GMTIRPLVDLLAVKKKQ---------ETKRSINEEIHTQFLDH-------LLTGIE---- 539

Qy     494 TALRMLLRTPTHTVHRYWR----KFDDSFMR 520
                     |  | :|:    :|:  :::
Db     540 -------DICGHYGHHHWKDKLNRFNKKYVK 563


RESULT 12
A57644
Na+/H+-exchanging protein 2 - human
C;Species: Homo sapiens (man)
C;Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 28-May-1999
C;Accession: A57644
R;Ghishan, F.K.; Knobel, S.M.; Summar, M.
Genomics 30, 25-30, 1995
A;Title: Molecular cloning, sequencing, chromosomal localization, and tissue distributic
A;Reference number: A57644; MUID:96129297
A;Accession: A57644
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-698 <GHI>
A;Cross-references: GB:S81591; NID:g1478393; PIDN:AAB36180.1; PID:g1478394
C;Genetics:
A;Gene: GDB:SLC9A2
A;Cross-references: GDB:132596; OMIM:600530
A;Map position: 2q11.2-2q11.2
C;Keywords: transmembrane protein

  Query Match             16.8%;  Score 466;  DB 2;  Length 698;
  Best Local Similarity   28.4%;  Pred. No. 2.1e-27;
  Matches  137;  Conservative   77;  Mismatches  194;  Indels   74;  Gaps   16;

Qy      56 IIGLCTGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLF 115
              ::||  | :|    |           |:||:|||||: :||: :  : || |  ||  :
Db       1 MVGLLLGGIIF---GVDEKSPPAMKTDVFFLYLLPPIVLDAGYFMPTRPFFENLGTIFWY 57

Qy     116 GAIGTL-----ISCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQ- 169
             :|||     |  |:     :: |   ||  |     |  |:: :| | |  | |
Db      58 AVVGTLWNSIGIGLSLFGICQIEAFGLSDITLL---QNLLFGSLISAVDPVAVLAVFENI 114

Qy     170 DETPLLYSLVFGEGVVNDATSVVLFNAIQSF----DMTSFDPKIGLHFIGNFLYLFLSST 225
                || ||||| ::||| :|||:|  :||     : : |   |   | ||  : :
Db     115 HVNEQLYILVFGESLLNDAVTVVLYNLFKSFCQMKTIQTVDVFAG---IANFFVVGIGGV 171

Qy     226 FLGVGIGLLCAYIIKKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVM 285
           :|: :| : |:  :   |  :    |   : | |||||| ||:|:|||||: :  | : |
Db     172 LIGILLGFIAAFTTR---FTHNIRVIEPLFVFLYSYLSYITAEMFHLSGIMAITACAMTM 228

Qy     286 SHYTWHNVTESSRVTTRHSFATLSFVAETFIFLYVGMDALDIE---KWKFVKNSQGLSVA 342
           : |   ||:: |  | ::    || |:|| ||:::|:  :        | ||
Db     229 NKYVEENVSQKSYTTIKYFMKMLSSVSETLIFIFMGVSTVGKNHEWNWAFV--------- 279

Qy     343 VSSILVGLILVGRAAFVFPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYN 402
                :   |: ||  || |: : |  :   :  ::|: | || : || |||:  || :
Db     280 --CFTLAFCLIWRALGVFVLTQVINWFR---TIPLTFKDQFIIAYGGL-RGAICFALVF- 332

Qy     403 KFTTSGHTSLHENAIMITSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSE 462
                   |      : ||: : |: |:  : |:  :||:  |        |       ||
Db     333 ---LLPATVFPRKKLFITAAIVVIFFTVFILGITIRPLVEFL----DVKRSNKKQQAVSE 385

Qy     463 PSSPKHFTVPLLDNQPDSESDMITGPEVARPTALRMLLRTPTHTVHRYWR----KFDDSF 518
                   |     |:       |:                       |  | :||    |||| :
Db     386 EIHWCRF----FDHVKTGIEDVC-----------------GHWGHNFWRDKFKKFDDKY 423

Qy     519 MR 520
           :|
Db     424 LR 425


RESULT 13
A48858
Na+/H+-exchanging protein - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 01-Dec-2000
C;Accession: A48858; I46613
R;Reilly, R.F.; Hildebrandt, F.; Biemesderfer, D.; Sardet, C.; Pouyssegur, J.; Aronso
Am. J. Physiol. 261, F1088-F1094, 1991
A;Title: cDNA cloning and immunolocalization of a Na(+)-H+ exchanger in LLC-PK1 renal
A;Reference number: A48858; MUID:92087905
A;Accession: A48858
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-818 <REI>
A;Cross-references: GB:S71135; NID:g240706; PIDN:AAB20633.1; PID:g240707
A;Experimental source: LLC-PK1 kidney cell line
A;Note: sequence extracted from NCBI backbone (NCBIN:71135, NCBIP:71136)
R;Reilly, R.F.
Am. J. Physiol. 261, 1088-1094, 1991
A;Title: cDNA cloning and immunolocalization of a Na+-H+ exchanger in LLC-PK1 renal e
A;Reference number: I46613
A;Accession: I46613
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-682,'H',684-818 <RE2>
A;Cross-references: GB:M89631; NID:g164595; PIDN:AAA31092.1; PID:g164596

  Query Match             16.6%;  Score 460.5;  DB 2;  Length 818;
  Best Local Similarity   28.0%;  Pred. No. 6.5e-27;
  Matches  145;  Conservative  108;  Mismatches  181;  Indels   83;  Gaps   22;

Qy      25 MNLFVALLCACIVLGHLLEE-NRWVNESITALIIGLCTGVVILLLSGGKSSHLLVFSEDL 83
           ::|:: | |   :  |::   :  | ||   :::||  | :|  :  |::   |    ::
Db     105 ISLWILLACLMKIGFHVIPTISSIVPESCLLIVVGLLVGGLIKAV--GETPPFL--QSEV 160

Qy      84 FFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIIS--FGAVKIFKHLDID 141
           ||::|||||| :||: :  :||  |  ||:::|  :||| :   :     || :     |:
Db     161 FFLFLLPPIILDAGYFLPLRQFTENLGTILIFAVVGTLWNAFFLGGLMYAVCLVGGEQIN 220

Qy     142 FLDFGDYLAIGAIFAATDSVCTLQVLSQDE-TPLLYSLVFGEGVVNDATSVVLFNAIQSF 200
           :   | |  |:| :| | |  | |  :     ||: ||||| ::||| :|||::  : |
Db     221 NIGLLDNLLFGSIISAVDPVAVLAVFEEIHINELLHILVFGESLLNDAVTVVLYHLFEEF 280

Qy     201 DMTSFDPKIG-----LHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDREVAL 255
              ::| ::|      | |: :|  : |   |:||  |:: |:  :   |  |    |
Db     281 --ANYD-RVGIVDIVLGFL-SFFVVSLGGVFVGVVYGVIAAFTSR---FTSHIRVIEPLF 333

Qy     256 MMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETF 315
           : | ||::|: ||||:|||| :   |:||  |   |::  |  | ::     | |:||
Db     334 VFLYSYMAYLSAELFHLSGIMALIASGVVMRPYVEANISHKSHTTIKYFLKMWSSVSETL 393
```

```
Qy    316 IFLYVGMDALDIE---KWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKKN 372
          ||:::|:  :       | ||         :|::|    |: |   |  |::  |   |
Db    394 IFIFLGVSTVAGSHHWNWTFV---------ISTLL--FCLIARVLGVLGLTWFIN---KF 439

Qy    373 SSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHEN-----AIMITSTVTVVL 427
             |:: : | || : || |||:: :|        ||  | :|      : :|: :||:
Db    440 RIVKLTPKDQFIIAYGGL-RGAIAFSL--------GHL-LDKNHFPMCDLFLTAIITVIF 489

Qy    428 FSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMITG 487
          |:  | |:  :||::||    ||          :: |  :      ||:       ::||
Db    490 FTVFVQGMTIRPLVDLLAVKKKQ---------ETKRSINEEIHTQFLDH-------LLTG 533

Qy    488 PEVARPTALRMLLRTPTHTVHRYWR----KFDDSFMR 520
          |              |  | :|:    :|:  :::
Db    534 IE-----------DICGHYGHHHWKDKLNRFNKKYVK 559


RESULT 14
A40204
Na+/H+-exchanging protein 1 - rat
N;Alternate names: Na+/H+ antiporter
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 28-Feb-1997
C;Accession: A40204
R;Orlowski, J.; Kandasamy, R.A.; Shull, G.E.
J. Biol. Chem. 267, 9331-9339, 1992
A;Title: Molecular cloning of putative members of the Na/H exchanger gene family. cDNA c
rally related proteins.
A;Reference number: A40204; MUID:92250539
A;Accession: A40204
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-820 <ORL>
A;Cross-references: GB:M85299
C;Keywords: transmembrane protein

  Query Match             16.6%;  Score 460;  DB 2;  Length 820;
  Best Local Similarity   28.5%;  Pred. No. 7.1e-27;
  Matches  139;  Conservative   98;  Mismatches  180;  Indels   70;  Gaps   19;

Qy     48 VNESITALIIGLCTGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFV 107
          | ||   :::||  |   |:   |::   |    |:||::|||||| :||: :  :||
Db    133 VPESCLLIVVGLLVGG--LIKGVGETPPFL--QSDVFFLFLLPPIILDAGYFLPLRQFTE 188

Qy    108 NFMTIMLFGAIGTLISCSIIS--FGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQ 165
          |  ||::|  :||| :   :     || :     |: :   | |  |:| :| | |  |
Db    189 NLGTILIFAVVGTLWNAFFLGGLLYAVCLVGGEQINNIGLLDTLLFGSIISAVDPVAVLA 248

Qy    166 VLSQDE-TPLLYSLVFGEGVVNDATSVVLFNAIQSFDMTSFDPKIGLH--FIG--NFLYL 220
          | :     ||: |||| ::||| :|||::  : |    :   :|:   |:|  :|  :
Db    249 VFEEIHINELLHILVFGESLLNDAVTVVLYHLFEEFASYEY---VGISDIFLGFLSFFVV 305

Qy    221 FLSSTFLGVGIGLLCAYIIKKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFF 280
          |   |:||  |:: |:  :   |  |     |   : | ||::|: ||||:|||||: :
Db    306 SLGGVFVGVVYGVIAAFTSR---FTSHIRVIEPLFVFLYSYMAYLSAELFHLSGIMALIA 362

Qy    281 CGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETFIFLYVGMDALDIE---KWKFVKNSQ 337
          |:||  |   |::  |  | ::     | |:|| ||:::|:  :          | ||
Db    363 SGVVMRPYVEANISHKSHTTIKYFLKMWSSVSETLIFIFLGVSTVAGSHQWNWTFV---- 418

Qy    338 GLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSI 397
             :|:::|    |: |   |  |::  |   |    |:: : | || : || |||::
Db    419 -----ISTLL--FCLIARVLGVLVLTWFIN---KFRIVKLTPKDQFIIAYGGL-RGAIAF 467

Qy    398 ALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSS 457
          :| |       |  : :  : :|: :||: |:  | |:  :||::||    ||
Db    468 SLGY--LLDKKHFPMCD--LFLTAIITVIFFTVFVQGMTIRPLVDLLAVKKKQ------- 516

Qy    458 MTTSEPSSPKHFTVPLLDNQPDSESDMITGPEVARPTALRMLLRTPTHTVHRYWR----K 513
             :: |  :      ||:       ::|| |                  | | :|:    :
Db    517 --ETKRSINEEIHTQFLDH-------LLTGIE-----------DICGHYGHHHWKDKLNR 556

Qy    514 FDDSFMR 520

          |:  :::
Db    557 FNKKYVK 563


RESULT 15
S16328
Na+/H+-exchanging protein - rabbit
N;Alternate names: Na+/H+ antiporter; Na+/H+ exchanger; pH regulatory protein
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 01-Dec-2000
C;Accession: S16328; S13926; S30602
R;Tse, C.M.; Ma, A.I.; Yang, V.W.; Watson, A.J.M.; Levine, S.; Montrose, M.H.; Potter
EMBO J. 10, 1957-1967, 1991
A;Title: Molecular cloning and expression of a cDNA encoding the rabbit ileal villus
A;Reference number: S16328; MUID:91293066
A;Accession: S16328
A;Molecule type: mRNA
A;Residues: 1-816 <TSE>
A;Cross-references: EMBL:X59935; NID:g1642; PIDN:CAA42558.1; PID:g1643
R;Fliegel, L.; Sardet, C.; Pouyssegur, J.; Barr, A.
FEBS Lett. 279, 25-29, 1991
A;Title: Identification of the protein and cDNA of the cardiac Na(+)/H(+) exchanger.
A;Reference number: S13926; MUID:91138752
A;Accession: S13926
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 472-816 <FLI>
A;Cross-references: EMBL:X56536; NID:g1666; PIDN:CAA39881.1; PID:g1667
R;Hildebrandt, F.; Pizzonia, J.H.; Reilly, R.F.; Reboucas, N.A.; Sardet, C.; Pouysseg
Biochim. Biophys. Acta 1129, 105-108, 1991
A;Title: Cloning, sequence, and tissue distribution of a rabbit renal Na(+)/H(+) exch
A;Reference number: S30602; MUID:92096447
A;Accession: S30602
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-241,'A',243-568,'E',570-816 <HIL>
A;Cross-references: EMBL:X61504; NID:g1653; PIDN:CAA43721.1; PID:g1654
A;Note: the authors translated the codon CTG for residue 646 as Val and GCG for resid
C;Keywords: glycoprotein; transmembrane protein
F;16-35/Domain: transmembrane #status predicted <TM1>
F;108-127/Domain: transmembrane #status predicted <TM2>
F;130-149/Domain: transmembrane #status predicted <TM3>
F;155-174/Domain: transmembrane #status predicted <TM4>
F;192-211/Domain: transmembrane #status predicted <TM5>
F;295-315/Domain: transmembrane #status predicted <TM6>
F;339-357/Domain: transmembrane #status predicted <TM7>
F;387-406/Domain: transmembrane #status predicted <TM8>
F;411-430/Domain: transmembrane #status predicted <TM9>
F;481-500/Domain: transmembrane #status predicted <TM10>
F;75,370/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             16.6%;  Score 459.5;  DB 2;  Length 816;
  Best Local Similarity   28.2%;  Pred. No. 7.7e-27;
  Matches  144;  Conservative  105;  Mismatches  191;  Indels   71;  Gaps   20;

Qy     25 MNLFVALLCACIVLGHLLEE-NRWVNESITALIIGLCTGVVILLLSGGKSSHLLVFSEDL 83
          ::|:: | |   :  |::   :  | ||   :::||  | :|  : | |   |     ::
Db    105 ISLWILLACLMKIGFHVIPTISSIVPESCLLIVVGLLVGGLIKGV-GEKPPFL---QSEV 160

Qy     84 FFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIIS--FGAVKIFKHLDID 141
          ||::|||||| :||: :  :||  |  ||::|  :||| :   :     || :     |:
Db    161 FFLFLLPPIILDAGYFLPLRQFTENLGTILIFAVVGTLWNAFFLGGLMYAVCLVGGEQIN 220

Qy    142 FLDFGDYLAIGAIFAATDSVCTLQVLSQDE-TPLLYSLVFGEGVVNDATSVVLFNAIQSF 200
          :   | |  |:| :| | |  | |  :     ||: ||||| ::||| :|||::  : |
Db    221 NIGLLDNLLFGSIISAVDPVAVLAVFEEIHINELLHILVFGESLLNDAVTVVLYHLFEEF 280

Qy    201 DMTSFD----PKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDREVALM 256
             ::|       | | |: :|  : |   |:||  |:: |:  :   |  |     |  :
Db    281 --ANYDHVGIVDIVLGFL-SFFVVALGGVFVGVVYGVIAAFTSR---FTAHIRVIEPLFV 334

Qy    257 MLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETFI 316
          | ||::|: ||||:|||||: :   |:||  |   |::  |  | ::     | |:|| |
```

```
Db    335 FLYSYMAYLSAELFHLSGIMALIASGVVMRPYVEANISHKSHTTIKYFLKMWSSVSETLI 394

Qy    317 FLYVGMDALDIE---KWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKKNS 373
          |:::|:  :       | ||          :|::|    |: |   |  |::  |   |
Db    395 FIFLGVSTVAGSHHWNWTFV---------ISTLL--FCLIARVLGVLGLTWFIN---KFR 440

Qy    374 SDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTVVF 433
            |:: : | || : || |||:: :| |       |  : :  : :|: :||: |:  |
Db    441 IVKLTPKDQFIIAYGGL-RGAIAFSLGY--LLDKKHFPMCD--LFLTAIITVIFFTVFVQ 495

Qy    434 GLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMITGPEVARP 493
          |:  :||::||    ||          :: |  :      ||:       ::|| |
Db    496 GMTIRPLVDLLAVKKKQ---------ETKRSINEEIHTQFLDH-------LLTGIE---- 535

Qy    494 TALRMLLRTPTHTVHRYWR----KFDDSFMR 520
                    |  | :|:    :|:  :::
Db    536 -------DICGHYGHHHWKDKLNRFNKKYVK 559
```

Search completed: October 18, 2002, 12:25:49
Job time : 26 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: October 18, 2002, 12:24:11 ; Search time 13 Seconds
(without alignments)
1614.307 Million cell updates/sec

Title: US-09-830-123-2
Perfect score: 2768
Sequence: 1 MAFGLSSLLQNSDLFTSDHA.....FGGRGFVPFVAGSPVEQSPR 542

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	578.5	20.9	669	1	NAH6_HUMAN	Q92581	homo sapien
2	568	20.5	633	1	NAH2_YEAST	Q04121	saccharomyc
3	485.5	17.5	813	1	NAH2_RAT	P48763	rattus norv
4	482.5	17.4	809	1	NAH2_RABIT	P50482	oryctolagus
5	481.5	17.4	812	1	NAH2_HUMAN	Q9uby0	homo sapien
6	472.5	17.1	820	1	NAH1_MOUSE	Q61165	mus musculu
7	471.5	17.0	832	1	NAH3_RABIT	P26432	oryctolagus
8	468.5	16.9	831	1	NAH3_RAT	P26433	rattus norv
9	468	16.9	759	1	NAHB_ONCMY	Q01345	oncorhynchu
10	467.5	16.9	820	1	NAH1_RAT	P26431	rattus norv
11	467.5	16.9	822	1	NAH1_CRIGR	P48761	cricetulus
12	466	16.8	839	1	NAH3_DIDMA	Q28362	didelphis m
13	461	16.7	834	1	NAH3_HUMAN	P48764	homo sapien
14	460.5	16.6	818	1	NAH1_PIG	P48762	sus scrofa
15	459.5	16.6	816	1	NAH1_RABIT	P23791	oryctolagus
16	458.5	16.6	817	1	NAH1_BOVIN	Q28036	bos taurus
17	456.5	16.5	815	1	NAH1_HUMAN	P19634	homo sapien
18	437.5	15.8	717	1	NAH4_RAT	P26434	rattus norv
19	425	15.4	896	1	NAH5_HUMAN	Q14940	homo sapien
20	412.5	14.9	898	1	NAH5_RAT	Q9z0x2	rattus norv
21	335.5	12.1	609	1	NAH_CAEEL	P35449	caenorhabdi
22	217.5	7.9	542	1	YM87_MYCTU	Q50678	mycobacteri
23	211.5	7.6	549	1	YJCE_ECOLI	P32703	escherichia
24	191	6.9	578	1	YCGO_ECOLI	P76007	escherichia
25	169	6.1	759	1	NAH2_SCHPO	O14123	schizosacch
26	166.5	6.0	808	1	NAH2_ZYGRO	O42701	zygosacchar
27	156.5	5.7	426	1	Y057_METJA	Q60362	methanococc
28	143.5	5.2	383	1	NAPA_ENTHR	P26235	enterococcu
29	134	4.8	791	1	NAH1_ZYGRO	Q99173	zygosacchar
30	130	4.7	422	1	YF21_METJA	Q58916	methanococc
31	124	4.5	519	1	NU4M_PODAN	P15582	podospora a
32	123.5	4.5	400	1	TYRP_HAEIN	P44727	haemophilus
33	122.5	4.4	554	1	NU5M_APILI	P34855	apis mellif
34	122.5	4.4	558	1	YBAL_ECOLI	P39830	escherichia
35	122	4.4	512	1	NUOM_RHOCA	P50974	rhodobacter
36	122	4.4	574	1	RFT1_YEAST	P38206	saccharomyc
37	121	4.4	383	1	YJK8_YEAST	P42946	saccharomyc
38	119.5	4.3	468	1	NAH_SCHPO	P36606	schizosacch
39	119.5	4.3	488	1	NU4M_ASPAM	P03913	aspergillus
40	119.5	4.3	549	1	COX1_LEITA	P14544	leishmania
41	117.5	4.2	633	1	Y147_HAEIN	P44543	haemophilus
42	116.5	4.2	407	1	GLUP_HELPY	O25788	helicobacte
43	116.5	4.2	502	1	PUTP_ECOLI	P07117	escherichia
44	116.5	4.2	654	1	NU5M_RHIST	P50367	rhizopus st
45	116	4.2	430	1	YQGE_BACSU	P54487	bacillus su

ALIGNMENTS

RESULT 1
NAH6_HUMAN
ID NAH6_HUMAN STANDARD; PRT; 669 AA.
AC Q92581;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sodium/hydrogen exchanger 6 (Na(+)/H(+) exchanger 6) (NHE-6).
GN SLC9A6 OR NHE6 OR KIAA0267.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98175963; PubMed=9507001;
RA Numata M., Petrecca K., Lake N., Orlowski J.;
RT "Identification of a mitochondrial Na+/H+ exchanger.";
RL J. Biol. Chem. 273:6951-6959(1998).
RN [2]
RP SEQUENCE OF 4-669 FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=97191544; PubMed=9039502;
RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayasi Y.,
RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. VI.
RT The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
RT analysis of cDNA clones from cell line KG-1 and brain.";
RL DNA Res. 3:321-329(1996).
CC -!- FUNCTION: ELECTRONEUTRAL EXCHANGE OF PROTONS FOR NA+ AND K+ ACROSS
CC THE MITOCHONDRIAL INNER MEMBRANE. CONTRIBUTES TO ORGANELLAR VOLUME
CC AND CALCIUM HOMEOSTASIS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial.
CC -!- TISSUE SPECIFICITY: UBIQUITOUS; BUT IS MOST ABUNDANT IN
CC MITOCHONDRION-RICH TISSUES SUCH AS BRAIN, SKELETAL MUSCLE, AND
CC HEART.
CC -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC ---

CC ---
DR EMBL; AF030409; AAC39643.1; -.
DR EMBL; D87743; BAA13449.1; -.
DR MIM; 300231; -.
DR InterPro; IPR000676; NaH_Exchngr.
DR Pfam; PF00999; Na_H_Exchanger; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
KW Transmembrane; Sodium transport; Transport; Symport; Mitochondrion.
FT TRANSMEM 28 48 POTENTIAL.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 103 123 POTENTIAL.

```
FT   TRANSMEM    142    162       POTENTIAL.
FT   TRANSMEM    179    199       POTENTIAL.
FT   TRANSMEM    220    240       POTENTIAL.
FT   TRANSMEM    246    266       POTENTIAL.
FT   TRANSMEM    292    312       POTENTIAL.
FT   TRANSMEM    340    360       POTENTIAL.
FT   TRANSMEM    382    402       POTENTIAL.
FT   TRANSMEM    404    424       POTENTIAL.
FT   TRANSMEM    447    467       POTENTIAL.
FT   TRANSMEM    483    503       POTENTIAL.
SQ   SEQUENCE   669 AA;  74161 MW;  F6416596229F2639 CRC64;

  Query Match             20.9%;  Score 578.5;  DB 1;  Length 669;
  Best Local Similarity   29.6%;  Pred. No. 7.5e-30;
  Matches  165;  Conservative  110;  Mismatches  178;  Indels  105;  Gaps   19;

Qy      10 QNSDLFTSDHASVVSMNLFVALLCACIVLGHLLEEN--RWVNESITALIIGLCTGVVILL 67
           |  :    | |:::   :|: ||   |:   | :    |:::|:  |:| ||  |:|:
Db      62 QAEESHRQDSANLL---IFILLLTLTILTIWLFKHRRARFLHETGLAMIYGLLVGLVLRY 118

Qy      68 ---------------LSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMT 111
                          :   :: |: |  ::||  |||||||  ||: :|:: || |  :
Db     119 GIHVPSDVNNVTLSCEVQSSPTTLLVTFDPEVFFNILLPPIIFYAGYSLKRRHFFRNLGS 178

Qy     112 IMLFGAIGTLISCSIIS---FGAVKIFK---HLDIDFLDFGDYLAIGAIFAATDSVCTLQ 165
           |: :  :|| ||| :|    :| | : |    |  ||  | | |  ||| :||| |  |
Db     179 ILAYAFLGTAISCFVIGSIMYGCVTLMKVTGQLAGDFY-FTDCLLFGAIVSATDPVTVLA 237

Qy     166 VLSQDETPL-LYSLVFGEGVVNDATSVVLFNAI----------QSFDMTSFDPKIGLHFI 214
           :  : :  : ||:|:||| |:||| ::|| ::|           :||:|:    ||:
Db     238 IFHELQVDVELYALLFGESVLNDAVAIVLSSSIVAYQPAGDNSHTFDVTAMFKSIGI--- 294

Qy     215 GNFLYLFLSSTFLGVGIGLLCAYIIK--KLYFGRHSTDREVALMMLMSYLSYIMAELFYL 272
             || :| .|  :|   |:: | : |  ||   |     |  |  |||: ::::|| :
Db     295 --FLGIFSGSFAMGAATGVVTALVTKFTKL---REFQLLETGLFFLMSWSTFLLAEAWGF 349

Qy     273 SGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETFIFLYVGMDALDIEKWKF 332
           :|:: | ||||  :|||::|::  |:  |:  |  |:|:|| ||| |:|:     :   |
Db     350 TGVVAVLFCGITQAHYTYNNLSTESQHRTKQLFELLNFLAENFIFSYMGLTLFTFQNHVF 409

Qy     333 VKNSQGLSVAVSSILVG---LILVGRAAFVFPLSFLSNLAKKNSSDKISFRQQIIIWWAG 389
                     : :||    | :|||| ::||| | || :::   ||    | :: :||
Db     410 ----------NPTFVVGAFVAIFLGRAANIYPLSLLLNLGRRS---KIGSNFQHMMMFAG 456

Qy     390 LMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTVVFGLMTKPLINLLLPPHK 449
           | |||:: |||         |: :   :| ::|: :| |:  |||  |  ::: |
Db     457 L-RGAMAFALAIR------DTATYARQMMFSTTLLIVFFTVWVFGGGTTAMLSCL----- 504

Qy     450 QMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMITGPEVARPTALRMLLRTPTHTVHR 509
                               |  | :     ||: : :  ||  | |      :  :  : |
Db     505 -----------------HIRVGV-----DSDQEHLGVPENERRTT-----KAESAWLFR 536

Qy     510 YWRKFDDSFMRPVFGGRG 527
           |  || ::::|:    |
Db     537 MWYNFDHNYLKPLLTHSG 554


RESULT 2
NAH2_YEAST
ID   NAH2_YEAST     STANDARD;      PRT;   633 AA.
AC   Q04121;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Mitochondrial sodium/hydrogen exchanger (Mitochondrial Na(+)/H(+)
DE   exchanger).
GN   NHA2 OR NHX1 OR YDR456W OR D9461.40.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E., Berno A.,
RA   Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
RA   Hunicke-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
RA   Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
RA   Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
RA   Winant A., Yelton M., Botstein D., Davis R.W.;
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98175963; PubMed=9507001;
RA   Numata M., Petrecca K., Lake N., Orlowski J.;
RT   "Identification of a mitochondrial Na+/H+ exchanger.";
RL   J. Biol. Chem. 273:6951-6959(1998).
CC   -!- FUNCTION: ELECTRONEUTRAL EXCHANGE OF PROTONS FOR NA+ AND K+ ACROSS
CC       THE MITOCHONDRIAL INNER MEMBRANE. CONTRIBUTES TO ORGANELLAR VOLUME
CC       AND CALCIUM HOMEOSTASIS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial.
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U33007; AAB64861.1; -.
DR   SGD; S0002864; NHX1.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
KW   Transmembrane; Sodium transport; Transport; Symport; Mitochondrion.
FT   TRANSMEM     62     82       POTENTIAL.
FT   TRANSMEM     86    106       POTENTIAL.
FT   TRANSMEM    118    138       POTENTIAL.
FT   TRANSMEM    154    174       POTENTIAL.
FT   TRANSMEM    177    197       POTENTIAL.
FT   TRANSMEM    218    238       POTENTIAL.
FT   TRANSMEM    259    279       POTENTIAL.
FT   TRANSMEM    307    327       POTENTIAL.
FT   TRANSMEM    353    373       POTENTIAL.
FT   TRANSMEM    377    397       POTENTIAL.
FT   TRANSMEM    432    452       POTENTIAL.
FT   TRANSMEM    458    478       POTENTIAL.
SQ   SEQUENCE   633 AA;  70147 MW;  9B771ABDE41CEB0A CRC64;

  Query Match             20.5%;  Score 568;  DB 1;  Length 633;
  Best Local Similarity   27.2%;  Pred. No. 3.3e-29;
  Matches  166;  Conservative  117;  Mismatches  203;  Indels  124;  Gaps   19;

Qy       8 LLQNSDLFTSDHASV----VSMN----------LFVALLCACIVLG---HLLEEN-RWV 48
           || : ||  ||        | :|           ||: ||     |    :| ::  | |
Db      30 LLPSPDLPGSDDPIAGDPDVDLNPVTEEMFSSWALFIMLLLLISALWSSYYLTQKRIRAV 89

Qy      49 NESITALIIGLCTGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVN 108
           :|:: ::  |:  |::| :  |      : |:   ||  ||||||  |:|::: :  || |
Db      90 HETVLSIFYGMVIGLIIRMSPGHYIQDTVTFNSSYFFNVLLPPIILNSGYELNQVNFFNN 149

Qy     109 FMTIMLFGAIGTLISCSIISFGAVKIFKHLDIDFLD--FGDYLAIGAIFAATDSVCTLQV 166
             ::|::|   || ||  :|    : |:  | :: :|  | | ::||  :||| |  | :
Db     150 MLSILIFAIPGTFISAVVIGI-ILYIWTFLGLESIDISFADAMSVGATLSATDPVTILSI 208

Qy     167 LSQDET-PLLYSLVFGEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSST 225
           :  :  | ||:::||| ::||| |:|:|   | |               | ||  |  |
Db     209 FNAYKVDPKLYTIIFGESLLNDAISIVMFETCQKFHGQPATFSSVFEGAGLFLMTFSVSL 268

Qy     226 FLGVGIGLLCAYIIKKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVM 285
           :|| ||:| | ::|  :  |:    |  |::|::| ||  :   ::|||::: |||| :
Db     269 LIGVLIGILVALLLKHTHIRRY-PQIESCLILLIAYESYFFSNGCHMSGIVSLLFCGITL 327

Qy     286 SHYTWHNVTESSRVTTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSS 345
           || ::|::  |::| :: |  |: ::| |||:|:|::     :  :    : | : |::
```

```
Db    328 KHYAYYNMSRRSQITIKYIFQLLARLSENFIFIYLGLELFTEVELVY----KPLLIIVAA 383

Qy    346 ILVGLILVGRAAFVFPLS------------FLSNLAKKNSS--DKISFRQQIIIWWAGL 390
          |    | | |    |||||              :| :  :| |  |:| :  |:: :||||
Db    384 I---SICVARWCAVFPLSQFVNWIYRVKTIRSMSGITGENISVPDEIPYNYQMMTFWAGL 440

Qy    391 MRGAVSIALAYN-----KFTTSGHTSLHENAIMITSTVTVVLFSTVVFGLMTKPLINLL- 444
           |||| :|||       |||            :: : : ||: : ::||  |  :: :|
Db    441 -RGAVGVALALGIQGEYKFT-----------LLATVLVVVVLTVIIFGGTTAGMLEVLN 487

Qy    445 ------------------LPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMI 485
                             |    : :| |  |   | |          :| ||   |
Db    488 IKTGCISEEDTSDDEFDIEAPRAINLLNGSSIQTDLGPYSD--------NNSPDISIDQF 539

Qy    486 --------------------------TGPEVARPTALRMLLRTPTHTV----HRYWRK 513
                                     | |  || :  :  ||    |:     ::::
Db    540 AVSSNKNLPNNISTTGGNTFGGLNETENTSPNPARSSMDKRNLRDKLGTIFNSDSQWFQN 599

Qy    514 FDDSFMRPVF 523
          ||:  ::|||
Db    600 FDEQVLKPVF 609
```

```
RESULT 3
NAH2_RAT
ID   NAH2_RAT       STANDARD;      PRT;   813 AA.
AC   P48763; Q16434;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Sodium/hydrogen exchanger 2 (Na(+)/H(+) exchanger 2) (NHE-2) (H7).
GN   SLC9A2 OR NHE2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (LONG ISOFORM).
RC   TISSUE=Stomach;
RX   MEDLINE=93280160; PubMed=7685026;
RA   Wang Z., Orlowski J., Shull G.E.;
RT   "Primary structure and functional expression of a novel
RT   gastrointestinal isoform of the rat Na/H exchanger.";
RL   J. Biol. Chem. 268:11925-11928(1993).
RN   [2]
RP   SEQUENCE FROM N.A. (SHORT ISOFORM).
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Small intestine;
RX   MEDLINE=93248205; PubMed=7683411;
RA   Collins J.F., Honda T., Knobel S., Bulus N.M., Conary J.,
RA   Dubois R., Ghishan F.K.;
RT   "Molecular cloning, sequencing, tissue distribution, and functional
RT   expression of a Na+/H+ exchanger (NHE-2).";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:3938-3942(1993).
RN   [3]
RP   SEQUENCE FROM N.A. (SHORT ISOFORM).
RC   TISSUE=Liver;
RX   MEDLINE=96129297; PubMed=8595899;
RA   Ghishan F.K., Knobel S.M., Summar M.;
RT   "Molecular cloning, sequencing, chromosomal localization, and tissue
RT   distribution of the human Na+/H+ exchanger (SLC9A2).";
RL   Genomics 30:25-30(1995).
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT. SEEMS TO PLAY AN IMPORTANT ROLE IN
CC       COLONIC SODIUM ABSORPTION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC       SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY IN SMALL INTESTINE, COLON, AND
CC       STOMACH, WITH MUCH LOWER LEVELS IN SKELETAL MUSCLE, KIDNEY, BRAIN,
CC       TESTIS, UTERUS, HEART, AND LUNG.
CC   -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC       DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC   -!- CAUTION: REF.3 SEQUENCE WAS ORIGINALLY THOUGHT TO ORIGINATE FROM
CC       HUMAN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L11236; AAA72350.1; -.
DR   EMBL; L11004; AAA75406.1; -.
DR   EMBL; S81591; AAB36180.1; -.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation; Alternative splicing.
FT   DOMAIN        1     13       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       14     34       A (M1) HYDROPHOBIC.
FT   DOMAIN       35     80       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       81    101       B (M2) HYDROPHOBIC.
FT   DOMAIN      102    107       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    108    128       C (M3) (POTENTIAL).
FT   DOMAIN      129    139       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    140    160       D (M4) (POTENTIAL).
FT   DOMAIN      161    169       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    170    190       E (M5) (POTENTIAL).
FT   DOMAIN      191    209       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    210    230       F (M5A) (POTENTIAL).
FT   DOMAIN      231    237       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    238    258       G (M5B) (POTENTIAL).
FT   DOMAIN      259    278       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    279    299       H (M6) (POTENTIAL).
FT   DOMAIN      300    308       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    309    329       I (M7) (POTENTIAL).
FT   DOMAIN      330    361       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    362    382       J (M8) (POTENTIAL).
FT   DOMAIN      383    392       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    393    413       K (M9) (POTENTIAL).
FT   DOMAIN      414    430       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      431    451       L, HYDROPHOBIC.
FT   DOMAIN      452    459       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    460    480       M13 (POTENTIAL).
FT   DOMAIN      481    813       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    351    351       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC      1    116       MISSING (IN SHORT ISOFORM).
FT   CONFLICT    504    504       H -> HW (IN REF. 3).
FT   CONFLICT    610    616       LYQIRQR -> SLSNPPA (IN REF. 3).
FT   CONFLICT    742    742       A -> P (IN REF. 3).
FT   CONFLICT    786    786       V -> G (IN REF. 2).
SQ   SEQUENCE   813 AA;  91402 MW;  29727267D7085845 CRC64;
```

```
  Query Match             17.5%;  Score 485.5;  DB 1;  Length 813;
  Best Local Similarity   28.0%;  Pred. No. 7.6e-24;
  Matches  148;  Conservative   86;  Mismatches  215;  Indels   79;  Gaps   18;

Qy     14 LFTSDHASV---VSMNLFVALLCACIVLGHLLEE-NRWVNESITALIIGLCTGVVILLLS 69
          :|| |:  |    : |:: |     :  ||  :    | ||   :::|| | :|
Db     71 VFTLDYPHVQIPFEITLWILLASLAKIGFHLYHKLPTIVPESCLLIMVGLLLGGIIF--- 127

Qy     70 GGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTL-----ISC 124
          |           |:||:||||||: :||: :  : || |  ||  :  :|||     |
Db    128 GVDEKSPPAMKTDVFFLYLLPPIVLDAGYFMPTRPFFENLGTIFWYAVVGTLWNSIGIGL 187

Qy    125 SIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQ-DETPLLYSLVFGEG 183
          |:     :: |   ||  |     |  |:: :| | |  | |         || |||||
Db    188 SLFGICQIEAFGLSDITLL---QNLLFGSLISAVDPVAVLAVFENIHVNEQLYILVFGES 244
```

```
Qy    184 VVNDATSVVLFNAIQSF----DMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYII 239
          ::||| :|||:|  :||     : : |   |   | ||  : :    :|: :| : |:
Db    245 LLNDAVTVVLYNLFKSFCQMKTIQTVDVFAG---IANFFVVGIGGVLIGILLGFIAAFTT 301

Qy    240 KKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRV 299
          :   |  :    |   : | |||||| ||:|:|||||: :  | : |: |   ||:: |
Db    302 R---FTHNIRVIEPLFVFLYSYLSYITAEMFHLSGIMAITACAMTMNKYVEENVSQKSYT 358

Qy    300 TTRHSFATLSFVAETFIFLYVGMDALDIE---KWKFVKNSQGLSVAVSSILVGLILVGRA 356
          | ::    || |:|| ||:::|:  :       | ||              :   |: ||
Db    359 TIKYFMKMLSSVSETLIFIFMGVSTVGKNHEWNWAFV----------CFTLAFCLIWRA 407

Qy    357 AFVFPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENA 416
            || |: : |  :   :  ::|: | || : || |||:  || :        |
Db    408 LGVFVLTQVINWFR---TIPLTFKDQFIIAYGGL-RGAICFALVF----LLPATVFPRKK 459

Qy    417 IMITSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDN 476
          : ||: : |: |:  : |:  :||:  |       |       ||    : |     |:
Db    460 LFITAAIVVIFFTVFILGITIRPLVEFL----DVKRSNKKQQAVSEEIHCRFF-----DH 510

Qy    477 QPDSESDMITGPEVARPTALRMLLRTPTHTVHRYWR----KFDDSFMR 520
                |:                    |  | :||    |||| ::|
Db    511 VKTGIEDVC-----------------GHWGHNFWRDKFKKFDDKYLR 540


RESULT 4
NAH2_RABIT
ID   NAH2_RABIT     STANDARD;      PRT;   809 AA.
AC   P50482;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Sodium/hydrogen exchanger 2 (Na(+)/H(+) exchanger 2) (NHE-2).
GN   SLC9A2 OR NHE2.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NEW ZEALAND WHITE;
RX   MEDLINE=93280159; PubMed=7685025;
RA   Tse C.-M., Levine S.A., Yun C.H., Montrose M.H., Little P.J.,
RA   Pouyssegur J., Donowitz M.;
RT   "Cloning and expression of a rabbit cDNA encoding a serum-activated
RT   ethylisopropylamiloride-resistant epithelial Na+/H+ exchanger isoform
RT   (NHE-2).";
RL   J. Biol. Chem. 268:11917-11924(1993).
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT. SEEMS TO PLAY AN IMPORTANT ROLE IN
CC       COLONIC SODIUM ABSORPTION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: HIGH LEVELS IN INTESTINE AND KIDNEY.
CC   -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC       DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L13733; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation.
FT   DOMAIN        1     16       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      17    37       A (M1) (POTENTIAL).
FT   DOMAIN        38    79       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      80    100      B (M2) (POTENTIAL).
FT   DOMAIN        101   106      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      107   127      C (M3) (POTENTIAL).
FT   DOMAIN        128   138      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      139   159      D (M4) (POTENTIAL).
FT   DOMAIN        160   168      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      169   189      E (M5) (POTENTIAL).
FT   DOMAIN        190   208      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      209   229      F (M5A) (POTENTIAL).
FT   DOMAIN        230   236      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      237   257      G (M5B) (POTENTIAL).
FT   DOMAIN        258   277      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      278   298      H (M6) (POTENTIAL).
FT   DOMAIN        299   307      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      308   328      I (M7) (POTENTIAL).
FT   DOMAIN        329   360      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      361   381      J (M8) (POTENTIAL).
FT   DOMAIN        382   391      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      392   412      K (M9) (POTENTIAL).
FT   DOMAIN        413   429      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      430   450      L (M10) (POTENTIAL).
FT   DOMAIN        451   458      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      459   479      M (M11) (POTENTIAL).
FT   DOMAIN        480   809      CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      350   350      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   809 AA;  90744 MW;  DBD00B45443D87A6 CRC64;

  Query Match             17.4%;  Score 482.5;  DB 1;  Length 809;
  Best Local Similarity   27.8%;  Pred. No. 1.2e-23;
  Matches  147;  Conservative   86;  Mismatches  216;  Indels   79;  Gaps   18;

Qy     14 LFTSDHASV---VSMNLFVALLCACIVLGHLLEE-NRWVNESITALIIGLCTGVVILLLS 69  -
          :|| |:  |     : |:: |    :  ||  :    | ||   :::||  | :|
Db     70 VFTLDYPHVQIPFEITLWILLASLAKIGFHLYHKLPTIVPESCLLIMVGLLLGGIIF--- 126

Qy     70 GGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTL-----ISC 124
          |           |:|||:||||||: :||: :  : || |  ||  :  :|||     |
Db    127 GVDEKSPPAMKTDVFFLYLLPPIVLDAGYFMPTRPFFENLGTIFWYAVVGTLWNSIGIGV 186

Qy    125 SIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQ-DETPLLYSLVFGEG 183
          |:     :: |   ||  |     |  |:: :| | |  | |         || |||||
Db    187 SLFGICQIEAFGLSDITLL---QNLLFGSLISAVDPVAVLAVFENIHVNEQLYILVFGES 243

Qy    184 VVNDATSVVLFNAIQSF----DMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYII 239
          ::||| :|||:|  :||     : : |   |   | ||  : :    :|: :| : |:
Db    244 LLNDAVTVVLYNLFKSFCQMKTIETIDVFAG---IANFFVVGIGGVLIGIFLGFIAAFTT 300

Qy    240 KKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRV 299
          :   |  :    |   : | |||||| ||:|:|||||: :  | : |: |   ||:: |
Db    301 R---FTHNIRVIEPLFVFLYSYLSYITAEMFHLSGIMAITACAMTMNKYVEENVSQKSYT 357

Qy    300 TTRHSFATLSFVAETFIFLYVGMDALDIE---KWKFVKNSQGLSVAVSSILVGLILVGRA 356
          | ::    || |:|| ||:::|:  :       | ||              :   |: ||
Db    358 TIKYFMKMLSSVSETLIFIFMGVSTVGKNHEWNWAFV----------CFTLAFCLIWRA 406

Qy    357 AFVFPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENA 416
            || |: : |  :   :  ::|: | || : || |||:  || :
Db    407 LGVFVLTRVINWFR---TIPLTFKDQFIIAYGGL-RGAICFALVF----LLPAAVFPRKK 458

Qy    417 IMITSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDN 476
          : ||: : |: |:  : |:  :||:  |       |       ||    : |     |:
Db    459 LFITAAIVVIFFTVFILGITIRPLVEFL----DVKRSNKKQQAVSEEIHCRFF-----DH 509

Qy    477 QPDSESDMITGPEVARPTALRMLLRTPTHTVHRYWR----KFDDSFMR 520
                |:                    |  | :||    |||| ::|
Db    510 VKTGIEDVC-----------------GHWGHNFWRDKFKKFDDKYLR 539


RESULT 5
```

```
NAH2_HUMAN
ID   NAH2_HUMAN     STANDARD;      PRT;   812 AA.
AC   Q9UBY0;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Sodium/hydrogen exchanger 2 (Na(+)/H(+) exchanger 2) (NHE-2).
GN   SLC9A2 OR NHE2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Colon;
RX   MEDLINE=99375108; PubMed=10444453;
RA   Malakooti J., Dahdal R.Y., Schmidt L., Layden T.J., Dudeja P.K.,
RA   Ramaswamy K.;
RT   "Molecular cloning, tissue distribution, and functional expression of
RT   the human Na(+)/H(+) exchanger NHE2.";
RL   Am. J. Physiol. 277:G383-G390(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Hou S., Wohldmann P.;
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT. SEEMS TO PLAY AN IMPORTANT ROLE IN
CC       COLONIC SODIUM ABSORPTION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN SKELETAL MUSCLE, COLON AND
CC       KIDNEY. LOWER LEVELS IN THE TESTIS, PROSTATE, OVARY, AND SMALL
CC       INTESTINE.
CC   -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC       DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF073299; AAD41635.1; -.
DR   EMBL; AC007239; AAF19248.1; -.
DR   MIM; 600530; -.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation.
FT   DOMAIN        1     13       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       14     33       A (M1) HYDROPHOBIC.
FT   DOMAIN       34     79       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       80    100       B (M2) HYDROPHOBIC.
FT   DOMAIN      101    106       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    107    127       C (M3) (POTENTIAL).
FT   DOMAIN      128    138       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    139    159       D (M4) (POTENTIAL).
FT   DOMAIN      160    168       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    169    189       E (M5) (POTENTIAL).
FT   DOMAIN      190    208       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    209    229       F (M5A) (POTENTIAL).
FT   DOMAIN      230    236       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    237    257       G (M5B) (POTENTIAL).
FT   DOMAIN      258    277       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    278    298       H (M6) (POTENTIAL).
FT   DOMAIN      299    307       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    308    328       I (M7) (POTENTIAL).
FT   DOMAIN      329    360       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    361    381       J (M8) (POTENTIAL).
FT   DOMAIN      382    391       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    392    412       K (M9) (POTENTIAL).
FT   DOMAIN      413    429       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      430    450       L, HYDROPHOBIC.
FT   DOMAIN      451    458       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    459    479       M13 (POTENTIAL).
FT   DOMAIN      480    812       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    350    350       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   812 AA;  91519 MW;  17EE177DC3830D0A CRC64;

  Query Match             17.4%;  Score 481.5;  DB 1;  Length 812;
  Best Local Similarity   27.9%;  Pred. No. 1.4e-23;
  Matches  150;  Conservative   87;  Mismatches  220;  Indels   81;  Gaps   19;

Qy       6 SSLLQNSDL--FTSDHASV---VSMNLFVALLCACIVLGHLLEE-NRWVNESITALIIGL 59
           ::| : | |  || |:  |     : |:: |     :  ||  :    | ||   :::||
Db      60 TTLFEESRLPVFTLDYPHVQIPFEITLWILLASLAKIGFHLYHKLPTIVPESCLLIMVGL 119

Qy      60 CTGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIG 119
             | :|    |           |:||:||||||: :||: :  : || |  ||  :  :|
Db     120 LLGGIIF---GVDEKSPPAMKTDVFFLYLLPPIVLDAGYFMPTRPFFENIGTIFWYAVVG 176

Qy     120 TL-----ISCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQ-DETP 173
           ||     |  |:     :: |   ||  |      |  |:: :| | |  | |
Db     177 TLWNSIGIGVSLFGICQIEAFGLSDITLL---QNLLFGSLISAVDPVAVLAVFENIHVNE 233

Qy     174 LLYSLVFGEGVVNDATSVVLFNAIQSF----DMTSFDPKIGLHFIGNFLYLFLSSTFLGV 229
           || ||||| ::||| :|||:|  :||     : : |   |   | ||  : :    :|:
Db     234 QLYILVFGESLLNDAVTVVLYNLFKSFCQMKTIETIDVFAG---IANFFVVGIGGVLIGI 290

Qy     230 GIGLLCAYIIKKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYT 289
           :| : |:  :   |  :    |   : | |||||| ||:|:||||: :  | : |: |
Db     291 FLGFIAAFTTR---FTHNIRVIEPLFVFLYSYLSYITAEMFHLSGIMAITACAMTMNKYV 347

Qy     290 WHNVTESSRVTTRHSFATLSFVAETFIFLYVGMDALDIE---KWKFVKNSQGLSVAVSSI 346
             ||:: |  | ::    || |:|| ||:::|:  :       | ||
Db     348 EENVSQKSYTTIKYFMKMLSSVSETLIFIFMGVSTVGKNHEWNWAFV----------CF 396

Qy     347 LVGLILVGRAAFVFPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTT 406
           :  |: ||  || |: : |   :  :  ::|: | || : || |||:  || :
Db     397 TLAFCLMWRALGVFVLTQVIN---RFRTIPLTFKDQFIIAYGGL-RGAICFALVF----L 448

Qy     407 SGHTSLHENAIMITSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSP 466
                 : ||: : |: |:  : |:  :||:  |        |       ||
Db     449 LPAAVFPRKKLFITAAIVVIFFTVFILGITIRPLVEFL----DVKRSNKKQQAVSE---- 500

Qy     467 KHFTVPLLDNQPDSESDMITGPEVARPTALRMLLRTPTHTVHRYWR----KFDDSFMR 520
             | |:      |:                    |  | :||    ||| ::|
Db     501 -EIYCRLFDHVKTGIEDVC----------------GHWGHNFWRDKFKKFDDKYLR 539


RESULT 6
NAH1_MOUSE
ID   NAH1_MOUSE     STANDARD;      PRT;   820 AA.
AC   Q61165;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Sodium/hydrogen exchanger 1 (Na(+)/H(+) exchanger 1) (NHE-1).
GN   SLC9A1 OR NHE1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C;
RA   Dewey M.J., Bowman L.H.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
```

```
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC       TRANSDUCTION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC       DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC   -!- CAUTION: HYDROPHOBIC DOMAINS A, B AND L ARE NOT BELIEVED TO BE
CC       TRANSMEMBRANAL, BUT ONLY MEMBRANE-ASSOCIATED.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U51112; AAA92976.1; -.
DR   MGD; MGI:102462; Slc9a1.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation.
FT   DOMAIN       1     12       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      13     32       A (M1) HYDROPHOBIC.
FT   DOMAIN      33    105       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN     106    127       B (M2) HYDROPHOBIC.
FT   DOMAIN     128    130       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   131    150       C (M3) (POTENTIAL).
FT   DOMAIN     151    162       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   163    183       D (M4) (POTENTIAL).
FT   DOMAIN     184    188       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   189    210       E (M5) (POTENTIAL).
FT   DOMAIN     211    230       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   231    251       F (M5A) (POTENTIAL).
FT   DOMAIN     252    260       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   261    282       G (M5B) (POTENTIAL).
FT   DOMAIN     283    301       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   302    322       H (M6) (POTENTIAL).
FT   DOMAIN     323    336       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   337    357       I (M7) (POTENTIAL).
FT   DOMAIN     358    388       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   389    410       J (M8) (POTENTIAL).
FT   DOMAIN     411    416       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   417    438       K (M9) (POTENTIAL).
FT   DOMAIN     439    452       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN     453    473       L, HYDROPHOBIC.
FT   DOMAIN     474    482       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   483    503       M (M10) (POTENTIAL).
FT   DOMAIN     504    820       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD   374    374       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   820 AA;  91467 MW;  0589C4D08DD348BE CRC64;

  Query Match             17.1%;  Score 472.5;  DB 1;  Length 820;
  Best Local Similarity   28.4%;  Pred. No. 5.2e-23;
  Matches  145;  Conservative  106;  Mismatches  189;  Indels   71;  Gaps   21;

Qy      25 MNLFVALLCACIVLGHLLEE-NRWVNESITALIIGLCTGVVILLLSGGKSSHLLVFSEDL 83
           ::|:: | |   :  |::   :  | ||   :::|| |   |:   |::   |    |:
Db     109 ISLWILLACLMKIGFHVIPTISSIVPESCLLIVVGLLVGG--LIKGVGETPPFL--QSDV 164

Qy      84 FFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIIS--FGAVKIFKHLDID 141
           ||::|||||| :||: :  :||  |  ||::|  :||| :   :     || :     |:
Db     165 FFLFLLPPIILDAGYFLPLRQFTENLGTILIFAVVGTLWNAFFLGGLLYAVCLVGGEQIN 224

Qy     142 FLDFGDYLAIGAIFAATDSVCTLQVLSQDE-TPLLYSLVFGEGVVNDATSVVLFNAIQSF 200
            :   | |  |:| :| | |  | |  :    ||: ||||| ::||| :|||::  : |
Db     225 NIGLLDTLLFGSIISAVDPVAVLAVFEEIHINELLHILVFGESLLNDAVTVVLYHLFEEF 284

Qy     201 DMTSFDPKIGLH--FIG--NFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDREVALM 256
```

```
           |:|  :|:   |:|  :|  : |   |:||  |:: |:  :   |  |    |   :
Db     285 --ASYD-SVGISDIFLGFLSFFVVALGGVFVGVVYGVIAAFTSR---FTSHIRVIEPLFV 338

Qy     257 MLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETFI 316
           | ||::|: ||||:||||: :   |:||  |   |::  |  | ::     | |:|| |
Db     339 FLYSYMAYLSAELFHLSGIMALIASGVVMRPYVEANISHKSHTTIKYFLKMWSSVSETLI 398

Qy     317 FLYVGMDALDIE---KWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKKNS 373
           |:::|:  :       | ||         :|::|    |: |   |  |::  |   |
Db     399 FIFLGVSTVAGSHQWNWTFV---------ISTLL--FCLIARVLGVLVLTWFIN---KFR 444

Qy     374 SDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTVVF 433
             |:: : | || : || |||:: :| |       |  : :  : :|: :||: |:  |
Db     445 IVKLTPKDQFIIAYGGL-RGAIAFSLGY--LLDKKHFPMCD--LFLTAIITVIFFTVFVQ 499

Qy     434 GLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMITGPEVARP 493
           |:  :||::||    ||          :: | :      ||:       ::|| |
Db     500 GMTIRPLVDLLAVKKKQ---------ETKRSINEEIHTQFLDH-------LLTGIE---- 539

Qy     494 TALRMLLRTPTHTVHRYWR----KFDDSFMR 520
                  |  | :|:    :|:  :::
Db     540 -------DICGHYGHHHWKDKLNRFNKKYVK 563


RESULT 7
NAH3_RABIT
ID   NAH3_RABIT     STANDARD;      PRT;   832 AA.
AC   P26432;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Sodium/hydrogen exchanger 3 (Na(+)/H(+) exchanger 3) (NHE-3).
GN   SLC9A3 OR NHE3.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NEW ZEALAND WHITE; TISSUE=Ileal villus, and Kidney cortex;
RX   MEDLINE=92250540; PubMed=1374392;
RA   Tse C.-M., Brant S.R., Walker S.S., Pouyssegur J., Donowitz M.;
RT   "Cloning and sequencing of a rabbit cDNA encoding an intestinal and
RT   kidney-specific Na+/H+ exchanger isoform (NHE-3).";
RL   J. Biol. Chem. 267:9340-9346(1992).
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC       TRANSDUCTION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: INTESTINAL AND KIDNEY SPECIFIC. MOST ABUNDANT
CC       IN KIDNEY CORTEX, FOLLOWED EQUALLY BY ILEUM AND ASCENDING COLON,
CC       THEN KIDNEY MEDULLA AND JEJUNUM. IS ABSENT FROM DUODENUM AND
CC       DESCENDING COLON.
CC   -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC       DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M87007; AAA31420.1; -.
DR   PIR; A40205; A40205.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
```

```
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation; Polymorphism.
FT   DOMAIN        1     11       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     12     27       M1 (POTENTIAL).
FT   DOMAIN       28     59       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     60     79       M2 (POTENTIAL).
FT   DOMAIN       80     81       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     82    101       M3 (POTENTIAL).
FT   DOMAIN      102    110       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    111    130       M4 (POTENTIAL).
FT   DOMAIN      131    134       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    135    154       M5 (POTENTIAL).
FT   DOMAIN      155    180       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    181    200       M5A (POTENTIAL).
FT   DOMAIN      201    209       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    210    229       M5B (POTENTIAL).
FT   DOMAIN      230    249       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    250    269       M6 (POTENTIAL).
FT   DOMAIN      270    298       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    299    319       M7 (POTENTIAL).
FT   DOMAIN      320    339       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    340    359       M8 (POTENTIAL).
FT   DOMAIN      360    366       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    367    385       M9 (POTENTIAL).
FT   DOMAIN      386    435       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    436    455       M10 (POTENTIAL).
FT   DOMAIN      456    832       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    325    325       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT     144    144       L -> P.
SQ   SEQUENCE   832 AA;  92748 MW;  8C8BB7C296CF8740 CRC64;

  Query Match             17.0%;  Score 471.5;  DB 1;  Length 832;
  Best Local Similarity   28.5%;  Pred. No. 6.1e-23;
  Matches  148;  Conservative   92;  Mismatches  197;  Indels   83;  Gaps   19;

Qy      22 VVSMNLFVALLCACIV-LGHLLEENRWVNESITALIIGLCTGVVILLLSGGKSSHLLVF- 79
           :::: : || |   :  | |  :    | ||   :::|| | ::|      : |:  |
Db      56 IIALWVLVASLAKIVFHLSH--KVTSVVPESALLIVLGLVLGGIVL-----AADHIASFT 108

Qy      80 -SEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIISFGAVKIF--- 135
            :  :|| ||||||: :||: :  : || |  :|:|:  :||: : :        :|
Db     109 LTPTVFFFYLLPPIVLDAGYFMPNRLFFSNLGSILLYAVVGTVWNAATTGLSLYGVFLSG 168

Qy     136 --KHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQ-DETPLLYSLVFGEGVVNDATSVV 192
             | |  |||   |  |:: || | |  | |  :     :|: :|||| ::||| :||
Db     169 IMGELKIGLLDF---LLFGSLIAAVDPVAVLAVFEEVHVNEVLFIIVFGESLLNDAVTVV 225

Qy     193 LFNAIQSFDMTSFDPKIGLHF---IGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHST 249
           |:|  |||     |   |:     | :|  : |  | :||    | : : :   | :|
Db     226 LYNVFQSFVTLGGDKVTGVDCVKGIVSFFVVSLGGTLVGVVFAFLLSLVTR---FTKHVR 282

Qy     250 DREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLS 309
             |   : ::|||||: :|:  || || : ||||    |   |::| |  | |::   |:
Db     283 VIEPGFVFIISYLSYLTSEMLSLSSILAITFCGICCQKYVKANISEQSATTVRYTMKMLA 342

Qy     310 FVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVG--RAAFVFPLSFLSN 367
             ||| ||:::|: |:|   | :           : :|: |: |   ||  |   ::| |
Db     343 SGAETIIFMFLGISAVDPLIWTW---------NTAFVLLTLLFVSVFRAIGVVLQTWLLN 393

Qy     368 LAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAY----NKFTTSGHTSLHENAIMITSTV 423
             :    : |    |::: : || ||||: ||       ||       : |  : :::|:
Db     394 RYRMVQLELI---DQVVMSYGGL-RGAVAFALVALLDGNK--------VKEKNLFVSTTI 441

Qy     424 TVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESD 483
             || |: :  ||  |||:  |    |  | |     :|    : |   |      |  :
Db     442 IVVFFTVIFQGLTIKPLVQWL----KVKRSEHREPKLNEKLHGRAFDHIL------SAIE 491

Qy     484 MITGPEVARPTALRMLLRTPTHTVHRY----WRKFDDSFM 519
            |:|                  | |    |  ||  |:
Db     492 DISG----------------QIGHNYLRDKWANFDRRFL 514


RESULT 8
```

```
NAH3_RAT
ID   NAH3_RAT       STANDARD;      PRT;   831 AA.
AC   P26433;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Sodium/hydrogen exchanger 3 (Na(+)/H(+) exchanger 3) (NHE-3).
GN   SLC9A3 OR NHE3.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Kidney;
RX   MEDLINE=92250539; PubMed=1577762;
RA   Orlowski J., Kandasamy R.A., Shull G.E.;
RT   "Molecular cloning of putative members of the Na/H exchanger gene
RT   family. cDNA cloning, deduced amino acid sequence, and mRNA tissue
RT   expression of the rat Na/H exchanger NHE-1 and two structurally
RT   related proteins.";
RL   J. Biol. Chem. 267:9331-9339(1992).
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC       TRANSDUCTION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: MOST ABUNDANT IN COLON AND SMALL INTESTINE,
CC       FOLLOWED BY KIDNEY AND STOMACH.
CC   -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC       DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M85300; AAA41702.1; -.
DR   PIR; B40204; B40204.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation.
FT   DOMAIN        1     11       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       12     25       A (M1) HYDROPHOBIC.
FT   DOMAIN       26     51       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       52     70       B (M2) HYDROPHOBIC.
FT   DOMAIN       71     76       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     77     96       C (M3) (POTENTIAL).
FT   DOMAIN       97    109       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    110    130       D (M4) (POTENTIAL).
FT   DOMAIN      131    136       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    137    157       E (M5) (POTENTIAL).
FT   DOMAIN      158    177       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    178    199       F (M5A) (POTENTIAL).
FT   DOMAIN      200    207       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    208    229       G (M5B) (POTENTIAL).
FT   DOMAIN      230    249       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    250    271       H (M6) (POTENTIAL).
FT   DOMAIN      272    287       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    288    306       I (M7) (POTENTIAL).
FT   DOMAIN      307    337       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    338    359       J (M8) (POTENTIAL).
FT   DOMAIN      360    366       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    367    387       K (M9) (POTENTIAL).
FT   DOMAIN      388    402       EXTRACELLULAR (POTENTIAL).
```

```
FT   DOMAIN       403    423       L, HYDROPHOBIC.
FT   DOMAIN       424    432       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     433    453       M (M10) (POTENTIAL).
FT   DOMAIN       454    831       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     323    323       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   831 AA;  93105 MW;  77A4BF10DFF99E3F CRC64;

  Query Match             16.9%;  Score 468.5;  DB 1;  Length 831;
  Best Local Similarity   27.7%;  Pred. No. 9.4e-23;
  Matches  146;  Conservative   89;  Mismatches  200;  Indels   93;  Gaps   18;

Qy      22 VVSMNLFVALLCACIV-LGHLLEENRWVNESITALIIGLCTGVVILLLSGGKSSHLLVF- 79
           :::: : || |   :  | |  :    | ||   :::|| | ::       : |:  |
Db      54 IIALWILVASLAKIVFHLSH--KVTSVVPESALLIVLGLVLGGIV-----WAADHIASFT 106

Qy      80 -SEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIISFGAVKIF--- 135
            :  ||| ||||||: :||: :  : || |  ||:|:  |||: : :         :|
Db     107 LTPTLFFFYLLPPIVLDAGYFMPNRLFFGNLGTILLYAVIGTIWNAATTGLSLYGVFLSG 166

Qy     136 --KHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQ-DETPLLYSLVFGEGVVNDATSVV 192
             | |  |||   |  |:: || | |  | |  :     :|: :|||| ::||| :||
Db     167 LMGELKIGLLDF---LLFGSLIAAVDPVAVLAVFEEVHVNEVLFIIVFGESLLNDAVTVV 223

Qy     193 LFNAIQSFDMTSFDPKIGLHF---IGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHST 249
           |:|  :||     |   |:     | :|  : |  | :||    | : : :   | :|
Db     224 LYNVFESFVTLGGDAVTGVDCVKGIVSFFVVSLGGTLVGVIFAFLLSLVTR---FTKHVR 280

Qy     250 DREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLS 309
             |   : ::|||||: :|:  || || : ||||    |   |::| |  | |::   |:
Db     281 IIEPGFVFVISYLSYLTSEMLSLSAILAITFCGICCQKYVKANISEQSATTVRYTMKMLA 340

Qy     310 FVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLA 369
           ||| ||:::|: |:|   | :       : |   : :  | | ||  |   ::: |
Db     341 SGAETIIFMFLGISAVDPVIWTW-------NTAFVLLTLVFISVYRAIGVVLQTWILNRY 393

Qy     370 KKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFS 429
           :   : |    |::: : || ||||: ||          : |  : :::|: || |:
Db     394 RMVQLETI---DQVVMSYGGL-RGAVAYALV----VLLDEKKVKEKNLFVSTTLIVVFFT 445

Qy     430 TVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPK----------HFTVPLLDNQP 478
           :  ||  |||:  |             :  ||   ||           |    : |
Db     446 VIFQGLTIKPLVQWL-----------KVKRSEQREPKLNEKLHGRAFDHILSAIED--- 490

Qy     479 DSESDMITGPEVARPTALRMLLRTPTHTVHRY----WRKFDDSFMRPV 522
                 |:|                  | |    |  ||  |:  |
Db     491 ------ISG----------------QIGHNYLRDKWSNFDRKFLSKV 515


RESULT 9
NAHB_ONCMY
ID   NAHB_ONCMY     STANDARD;      PRT;   759 AA.
AC   Q01345;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Na(+)/H(+) exchanger beta (Na(+)/H(+) antiporter) (Beta-NHE).
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Nucleated erythrocyte;
RX   MEDLINE=92357712; PubMed=1379718;
RA   Borgese F., Sardet C., Cappadoro M., Pouyssegur J., Motais R.;
RT   "Cloning and expression of a cAMP-activated Na+/H+ exchanger:
RT   evidence that the cytoplasmic domain mediates hormonal regulation.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:6765-6769(1992).
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT.
CC   -!- SUBCELLULAR LOCATION: BASOLATERAL MEMBRANE.
CC   -!- PTM: ACTIVATED BY CAMP, PROTEIN KINASE A AND PROTEIN KINASE C.
CC   -!- MISCELLANEOUS: INHIBITED BY AMILORIDE AND 5-AMINO-SUBSTITUTED
CC       DERIVATIVES AND ACTIVATED IN A COOPERATIVE FASHION BY
CC       INTRACELLULAR H+.
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M94581; AAA49549.1; -.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation.
FT   DOMAIN         1     14       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      15     34       M1 (POTENTIAL).
FT   DOMAIN        35     75       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      76     95       M2 (POTENTIAL).
FT   DOMAIN        96     97       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      98    117       M3 (POTENTIAL).
FT   DOMAIN       118    122       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     123    142       M4 (POTENTIAL).
FT   DOMAIN       143    149       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     150    169       M5 (POTENTIAL).
FT   DOMAIN       170    195       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     196    214       M5A (POTENTIAL).
FT   DOMAIN       215    225       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     226    244       M5B (POTENTIAL).
FT   DOMAIN       245    261       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     262    282       M6 (POTENTIAL).
FT   DOMAIN       283    311       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     312    330       M7 (POTENTIAL).
FT   DOMAIN       331    352       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     353    372       M8 (POTENTIAL).
FT   DOMAIN       373    376       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     377    398       M9 (POTENTIAL).
FT   DOMAIN       399    446       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     447    467       M10 (POTENTIAL).
FT   DOMAIN       468    759       CYTOPLASMIC (POTENTIAL).
FT   MOD_RES      641    641       PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT   MOD_RES      648    648       PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT   CARBOHYD      49     49       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     338    338       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   759 AA;  85173 MW;  D6D95442995AE251 CRC64;

  Query Match             16.9%;  Score 468;  DB 1;  Length 759;
  Best Local Similarity   28.2%;  Pred. No. 9.2e-23;
  Matches  148;  Conservative  100;  Mismatches  184;  Indels   92;  Gaps   22;

Qy      31 LLCACIVLG-HLLEE-NRWVNESITALIIGLCTGVVILLLSGGKSSHLLVFSEDLFFIYL 88
           ||    : || ||:   :  | ||   :::||  | :| ::  |:   :|     |||: |
Db      78 LLALLMKLGFHLIPRLSAVVPESCLLIVVGLLVGGLIKVI--GEEPPVL--DSQLFFLCL 133

Qy      89 LPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLIS-----------CSIISFGAVKIFKH 137
           ||||| :||: :  : |  |  ||::|  |||| :             | | | |
Db     134 LPPIILDAGYFLPIRPFTENVGTILVFAVIGTLWNAFFMGGLLYALCQIESVG------- 186

Qy     138 LDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDE-TPLLYSLVFGEGVVNDATSVVLFNA 196
             :  :|    |  |:| :| | |  | |  :     |:: ||||| ::||| :|||:|
Db     187 --LSGVDLLACLLFGSIVSAVDPVAVLAVFEEIHINELVHILVFGESLLNDAVTVVLYNL 244

Qy     197 IQSFD----MTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDRE 252
           : |      :|  |  :|   :  |  : |     :|   | | |:  :   |  |:   |
Db     245 FEEFSKVGTVTVLDVFLG---VVCFFVVSLGGVLVGAIYGFLAAFTSR---FTSHTRVIE 298

Qy     253 VALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVA 312
```

```
Db    299 PLFVFLYSYMAYLSSEMFHLSGIMALIACGVVMRPYVEANISHKSYTTIKYFLKMWSSVS 358

Qy    313 ETFIFLYVGMDAL---DIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLA 369

Db    359 ETLIFIFLGVSTVAGPHAWNWTFV---------ITTVI--LCLVSRVLGVIGLTFIIN-- 405

Qy    370 KKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFS 429

Db    406 -KFRIVKLTKKDQFIVAYGGL-RGAIAFSLGY--LLSNSHQMRN---LFLTAIITVIFFT 458

Qy    430 TVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMITGPE 489

Db    459 VFVQGMTIRPLVELLAVKKKK---------ESKPSINEEIHTEFLDH-------LLTGVE 502

Qy    490 VARPTALRMLLRTPTHTVHRYWR----KFDDSFM-RPVFGGRGF 528

Db    503 -----------GVCGHYGHYHWKEKLNRFNKTYVKRWLIAGENF 535


RESULT 10
NAH1_RAT
ID   NAH1_RAT       STANDARD;      PRT;   820 AA.
AC   P26431;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Sodium/hydrogen exchanger 1 (Na(+)/H(+) exchanger 1) (NHE-1).
GN   SLC9A1 OR NHE1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Heart;
RX   MEDLINE=92250539; PubMed=1577762;
RA   Orlowski J., Kandasamy R.A., Shull G.E.;
RT   "Molecular cloning of putative members of the Na/H exchanger gene
RT   family. cDNA cloning, deduced amino acid sequence, and mRNA tissue
RT   expression of the rat Na/H exchanger NHE-1 and two structurally
RT   related proteins.";
RL   J. Biol. Chem. 267:9331-9339(1992).
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC       TRANSDUCTION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: NOT TISSUE SPECIFIC.
CC   -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC       DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC   -!- CAUTION: HYDROPHOBIC DOMAINS A, B AND L ARE NOT BELIEVED TO BE
CC       TRANSMEMBRANAL, BUT ONLY MEMBRANE-ASSOCIATED.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M85299; AAA98479.1; -.
DR   PIR; A40204; A40204.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation.
FT   DOMAIN        1     12       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       13     32       A (M1) HYDROPHOBIC.
FT   DOMAIN       33    105       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      106    127       B (M2) HYDROPHOBIC.
FT   DOMAIN      128    130       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    131    150       C (M3) (POTENTIAL).
FT   DOMAIN      151    162       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    163    183       D (M4) (POTENTIAL).
FT   DOMAIN      184    188       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    189    210       E (M5) (POTENTIAL).
FT   DOMAIN      211    230       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    231    251       F (M5A) (POTENTIAL).
FT   DOMAIN      252    260       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    261    282       G (M5B) (POTENTIAL).
FT   DOMAIN      283    301       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    302    322       H (M6) (POTENTIAL).
FT   DOMAIN      323    336       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    337    357       I (M7) (POTENTIAL).
FT   DOMAIN      358    388       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    389    410       J (M8) (POTENTIAL).
FT   DOMAIN      411    416       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    417    438       K (M9) (POTENTIAL).
FT   DOMAIN      439    452       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      453    473       L, HYDROPHOBIC.
FT   DOMAIN      474    482       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    483    503       M (M10) (POTENTIAL).
FT   DOMAIN      504    820       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    374    374       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   820 AA;  91647 MW;  58398DE74A9642FB CRC64;

  Query Match             16.9%;  Score 467.5;  DB 1;  Length 820;
  Best Local Similarity   28.0%;  Pred. No. 1.1e-22;
  Matches  143;  Conservative  106;  Mismatches  191;  Indels   71;  Gaps   20;

Qy     25 MNLFVALLCACIVLGHLLEE-NRWVNESITALIIGLCTGVVILLLSGGKSSHLLVFSEDL 83

Db    109 ISLWILLACLMKIGFHVIPTISSIVPESCLLIVVGLLVGG--LIKGVGETPPFL--QSDV 164

Qy     84 FFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIIS--FGAVKIFKHLDID 141

Db    165 FFLFLLPPIILDAGYFLPLRQFTENLGTILIFAVVGTLWNAFFLGGLLYAVCLVGGEQIN 224

Qy    142 FLDFGDYLAIGAIFAATDSVCTLQVLSQDE-TPLLYSLVFGEGVVNDATSVVLFNAIQSF 200

Db    225 NIGLLDTLLFGSIISAVDPVAVLAVFEEIHINELLHILVFGESLLNDAVTVVLYHLFEEF 284

Qy    201 DMTSFDPKIGLH--FIG--NFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDREVALM 256

Db    285 ASYEY---VGISDIFLGFLSFFVVSLGGVFVGVVYGVIAAFTSR---FTSHIRVIEPLFV 338

Qy    257 MLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETFI 316

Db    339 FLYSYMAYLSAELFHLSGIMALIASGVVMRPYVEANISHKSHTTIKYFLKMWSSVSETLI 398

Qy    317 FLYVGMDALDIE---KWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKKNS 373

Db    399 FIFLGVSTVAGSHQWNWTFV---------ISTLL--FCLIARVLGVLVLTWFIN---KFR 444

Qy    374 SDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTVVF 433

Db    445 IVKLTPKDQFIIAYGGL-RGAIAFSLGY--LLDKKHFPMCD--LFLTAIITVIFFTVFVQ 499

Qy    434 GLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMITGPEVARP 493

Db    500 GMTIRPLVDLLAVKKKQ---------ETKRSINEEIHTQFLDH-------LLTGIE---- 539

Qy    494 TALRMLLRTPTHTVHRYWR----KFDDSFMR 520

Db    540 -------DICGHYGHHHWKDKLNRFNKKYVK 563


RESULT 11
NAH1_CRIGR
ID   NAH1_CRIGR     STANDARD;      PRT;   822 AA.
```

```
AC   P48761;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Sodium/hydrogen exchanger 1 (Na(+)/H(+) exchanger 1) (NHE-1).
GN   SLC9A1 OR NHE1.
OS   Cricetulus griseus (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10029;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93192332; PubMed=8383540;
RA   Counillon L., Pouyssegur J.;
RT   "Nucleotide sequence of the Chinese hamster Na+/H+ exchanger NHE1.";
RL   Biochim. Biophys. Acta 1172:343-345(1993).
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC       TRANSDUCTION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC       DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC   -!- CAUTION: HYDROPHOBIC DOMAINS A, B AND L ARE NOT BELIEVED TO BE
CC       TRANSMEMBRANAL, BUT ONLY MEMBRANE-ASSOCIATED.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X68970; CAA48771.1; -.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation.
FT   DOMAIN        1     12       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       13     32       A (M1) HYDROPHOBIC.
FT   DOMAIN       33    105       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      106    127       B (M2) HYDROPHOBIC.
FT   DOMAIN      128    130       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    131    150       C (M3) (POTENTIAL).
FT   DOMAIN      151    162       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    163    183       D (M4) (POTENTIAL).
FT   DOMAIN      184    188       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    189    210       E (M5) (POTENTIAL).
FT   DOMAIN      211    230       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    231    251       F (M5A) (POTENTIAL).
FT   DOMAIN      252    260       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    261    282       G (M5B) (POTENTIAL).
FT   DOMAIN      283    301       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    302    322       H (M6) (POTENTIAL).
FT   DOMAIN      323    336       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    337    357       I (M7) (POTENTIAL).
FT   DOMAIN      358    388       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    389    410       J (M8) (POTENTIAL).
FT   DOMAIN      411    416       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    417    438       K (M9) (POTENTIAL).
FT   DOMAIN      439    452       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      453    473       L, HYDROPHOBIC.
FT   DOMAIN      474    482       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    483    503       M (M10) (POTENTIAL).
FT   DOMAIN      504    822       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    374    374       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   822 AA;  92003 MW;  E97C1ACD4EB88DAA CRC64;
```

```
  Query Match             16.9%;  Score 467.5;  DB 1;  Length 822;
  Best Local Similarity   28.2%;  Pred. No. 1.1e-22;
  Matches  144;  Conservative  107;  Mismatches  189;  Indels   71;  Gaps   21;

Qy      25 MNLFVALLCACIVLGHLLEE-NRWVNESITALIIGLCTGVVILLLSGGKSSHLLVFSEDL 83
           ::|:: | |   :  |::   :  | ||   :::||  |   |:  |::   |    |:
Db     109 ISLWILLACLMKIGFHVIPTISSIVPESCLLIVVGLLVGG--LIKGVGETPPFL--QSDV 164

Qy      84 FFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIIS--FGAVKIFKHLDID 141
           ||::|||||| :||: :  :|| |  ||::|  :||| :   :     || :     |:
Db     165 FFLFLLPPIILDAGYFLPLRQFTENLGTILIFAVVGTLWNAFFLGGLLYAVCLVGGEQIN 224

Qy     142 FLDFGDYLAIGAIFAATDSVCTLQVLSQDE-TPLLYSLVFGEGVVNDATSVVLFNAIQSF 200
           :   | |  |:| :| | |  : |  :     ||: ||||| ::||| :|||::  : |
Db     225 NIGLLDTLLFGSIISAVDPVAVVAVFEEIHINELLHILVFGESLLNDAVTVVLYHLFEEF 284

Qy     201 DMTSFDPKIGLH--FIG--NFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDREVALM 256
             ::|  ||:   |:|  :|  : |   |:|| |:: |:  :   |  |     |   :
Db     285 --ANYD-SIGISDIFLGFLSFFVVALGGVFVGVVYGVIAAFTSR---FTSHIRVIEPLFV 338

Qy     257 MLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETFI 316
           | ||::|: ||||:|||| :   |:||  |   |::  |  | ::     | |:|| |
Db     339 FLYSYMAYLSAELFHLSGIMALIASGVVMRPYVEANISHKSHTTIKYFLKMWSSVSETLI 398

Qy     317 FLYVGMDALDIE---KWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKKNS 373
           |:::|:  :       | ||         :|::|    |: |   |  |::  |   |
Db     399 FIFLGVSTVAGSHQWNWTFV---------ISTLL--FCLIARVLGVLVLTWFIN---KFR 444

Qy     374 SDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTVVF 433
             |:: : | || : || |||:: :| |       |  : :  : :|: :||: |:  |
Db     445 IVKLTPKDQFIIAYGGL-RGAIAFSLGY--LMDKKHFPMCD--LFLTAIITVIFFTVFVQ 499

Qy     434 GLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMITGPEVARP 493
           |:  :||::||    ||          :: |  :      ||:       ::|| |
Db     500 GMTIRPLVDLLAVKKKQ---------ETKRSINEEIHTQFLDH-------LLTGIE---- 539

Qy     494 TALRMLLRTPTHTVHRYWR----KFDDSFMR 520
                 |  | :|:    :|:  :::
Db     540 -------DICGHYGHHHWKDKLNRFNKKYVK 563
```

```
RESULT 12
NAH3_DIDMA
ID   NAH3_DIDMA     STANDARD;      PRT;   839 AA.
AC   Q28362;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Sodium/hydrogen exchanger 3 (Na(+)/H(+) exchanger 3) (NHE-3).
GN   SLC9A3 OR NHE3.
OS   Didelphis marsupialis virginiana (North American opossum).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.
OX   NCBI_TaxID=9267;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95358256; PubMed=7631739;
RA   Amemiya M., Yamaji Y., Cano A., Moe O.W., Alpern R.J.;
RT   "Acid incubation increases NHE-3 mRNA abundance in OKP cells.";
RL   Am. J. Physiol. 269:C126-C133(1995).
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC       TRANSDUCTION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC       DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L42522; AAA98816.1; -.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation.
FT   DOMAIN        1     13       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     14     31       M1 (POTENTIAL).
FT   DOMAIN       32     65       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     66     85       M2 (POTENTIAL).
FT   DOMAIN       86     87       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     88    107       M3 (POTENTIAL).
FT   DOMAIN      108    113       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    114    133       M4 (POTENTIAL).
FT   DOMAIN      134    150       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    151    170       M5 (POTENTIAL).
FT   DOMAIN      171    186       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    187    206       M5A (POTENTIAL).
FT   DOMAIN      207    215       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    216    235       M5B (POTENTIAL).
FT   DOMAIN      236    255       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    256    276       M6 (POTENTIAL).
FT   DOMAIN      277    299       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    300    319       M7 (POTENTIAL).
FT   DOMAIN      320    347       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    348    366       M8 (POTENTIAL).
FT   DOMAIN      367    370       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    373    392       M9 (POTENTIAL).
FT   DOMAIN      393    442       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    443    462       M10 (POTENTIAL).
FT   DOMAIN      463    839       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    331    331       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   839 AA;  94765 MW;  F6F9EF778D5DBBB2 CRC64;

  Query Match             16.8%;  Score 466;  DB 1;  Length 839;
  Best Local Similarity   29.2%;  Pred. No. 1.4e-22;
  Matches  133;  Conservative   89;  Mismatches  186;  Indels   48;  Gaps   14;

Qy      22 VVSMNLFVALLCACIV-LGHLLEENRWVNESITALIIGLCTGVVILLLSGGKSSHLLVF- 79
           :::: : || |   :  | |  :    | ||   :::||  | ::       : |:  |
Db      62 IIALWILVASLAKIVFHLSH--KVTSVVPESALLIVLGLILGGIV-----WAADHIASFT 114

Qy      80 -SEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIISFGAVKIF--- 135
            :  :|| ||||||: :||: :  : || |  ||:|:  |||: : :         ::
Db     115 LTPTVFFFYLLPPIVLDAGYFMPNRLFFGNLGTILLYAVIGTVWNAATTGLSLYGVYLSG 174

Qy     136 --KHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQ-DETPLLYSLVFGEGVVNDATSVV 192
             | |  |||    |  |:: || | |  | |  :     :|: :|||| ::||| :||
Db     175 IMGDLSIGLLDF---LLFGSLIAAVDPVAVLAVFEEVHVNDVLFIIVFGESLLNDAVTVV 231

Qy     193 LFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKL-YFGRHSTDR 251
           |:|   ||      |   |:  :   :  |: |   |  ||:: |:::  :  | :|
Db     232 LYNVFDSFVSLGADKVTGVDCVKGIVSFFVVS-LGGTLIGIIFAFLLSLVTRFTKHVRII 290

Qy     252 EVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFV 311
           |   : ::||||||: :|:  || ||. : ||||    |   |::| |  | |::   |:
Db     291 EPGFVFIISYLSYLTSEMLSLSAILAITFCGICCQKYVKANISEQSATTVRYTMKMLASG 350

Qy     312 AETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKK 371
           ||| ||:::|: |:|   | :       : |   : :  | | ||  |   ::| |   |
Db     351 AETIIFMFLGISAVDPAIWTW-------NTAFILLTLVFISVYRAIGVVLQTWLLN---K 400

Qy     372 NSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTV 431
             ::    |::: : || ||||: ||          : |  : :::|: || |: :
Db     401 YRMVQLEIIDQVVMSYGGL-RGAVAYALV----VLLDEKKVKEKNLFVSTTIIVVFFTVI 455

Qy     432 VFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPK 467
           ||  |||:  |            :  ||   ||
Db     456 FQGLTIKPLVQWL-----------KVKKSEHREPK 479


RESULT 13
NAH3_HUMAN
ID   NAH3_HUMAN     STANDARD;      PRT;   834 AA.
AC   P48764;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Sodium/hydrogen exchanger 3 (Na(+)/H(+) exchanger 3) (NHE-3).
GN   SLC9A3 OR NHE3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney cortex;
RX   MEDLINE=95358265; PubMed=7631746;
RA   Brant S.R., Yun C.H., Donowitz M., Tse C.-M.;
RT   "Cloning, tissue distribution, and functional analysis of the human
RT   Na+/N+ exchanger isoform, NHE3.";
RL   Am. J. Physiol. 269:C198-C206(1995).
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC       TRANSDUCTION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC       DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U28043; AAB48990.1; -.
DR   MIM; 182307; -.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation.
FT   DOMAIN        1     12       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       13     24       A (M1) HYDROPHOBIC.
FT   DOMAIN       25     54       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       55     73       B (M2) HYDROPHOBIC.
FT   DOMAIN       74     79       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     80     99       C (M3) (POTENTIAL).
FT   DOMAIN      100    112       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    113    133       D (M4) (POTENTIAL).
FT   DOMAIN      134    139       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    140    160       E (M5) (POTENTIAL).
FT   DOMAIN      161    180       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    181    202       F (M5A) (POTENTIAL).
FT   DOMAIN      203    210       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    211    232       G (M5B) (POTENTIAL).
FT   DOMAIN      233    252       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    253    274       H (M6) (POTENTIAL).
FT   DOMAIN      275    290       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    291    309       I (M7) (POTENTIAL).
FT   DOMAIN      310    340       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    341    362       J (M8) (POTENTIAL).
FT   DOMAIN      363    369       CYTOPLASMIC (POTENTIAL).
```

```
FT   TRANSMEM     370    390       K (M9) (POTENTIAL).
FT   DOMAIN       391    405       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN       406    426       L, HYDROPHOBIC.
FT   DOMAIN       427    435       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     436    456       M (M10) (POTENTIAL).
FT   DOMAIN       457    834       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     241    241       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     326    326       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   834 AA;  92907 MW;  3E7CE33D65DF36F7 CRC64;

  Query Match             16.7%;  Score 461;  DB 1;  Length 834;
  Best Local Similarity   28.6%;  Pred. No. 2.8e-22;
  Matches  140;  Conservative   85;  Mismatches  197;  Indels   68;  Gaps   15;

Qy      48 VNESITALIIGLCTGVVILLLSGGKSSHLLVF--SEDLFFIYLLPPIIFNAGFQVKKKQF 105
           | ||   :::||  | ::       : |:  |  :  :|| ||||||: :||: :  : |
Db      82 VPESALLIVLGLVLGGIV-----WAADHIASFTLTPTVFFFYLLPPIVLDAGYFMPNRLF 136

Qy     106 FVNFMTIMLFGAIGTLISCSIISFGAVKIF-----KHLDIDFLDFGDYLAIGAIFAATDS 160
           | |  ||:|:  :||: : :        :|       | |  |||   |  |:: || |
Db     137 FGNLGTILLYAVVGTVWNAATTGLSLYGVFLSGLMGDLQIGLLDF---LLFGSLMAAVDP 193

Qy     161 VCTLQVLSQ-DETPLLYSLVFGEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHF---IGN 216
           |  | |  :    :|: :|||| ::||| :|||:|  :||     |   |:     | :
Db     194 VAVLAVFEEVHVNEVLFIIVFGESLLNDAVTVVLYNVFESFVALGGDNVTGVDCVKGIVS 253

Qy     217 FLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGIL 276
           |  : |  | :||    | : : :   | :|    |   : ::||||||: :|:  || ||
Db     254 FFVVSLGGTLVGVVFAFLLSLVTR---FTKHVRIIEPGFVFIISYLSYLTSEMLSLSAIL 310

Qy     277 TVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNS 336
           : ||||    |   |::| |  | |::   |:  ||| ||:::|: |::   | :
Db     311 AITFCGICCQKYVKANISEQSATTVRYTMKMLASSAETIIFMFLGISAVNPFIWTW---- 366

Qy     337 QGLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVS 396
           : |   : :  | | ||  |   ::| |  :    : |    |::: : || ||||:
Db     367 ---NTAFVLLTLVFISVYRAIGVVLQTWLLNRYRMVQLEPI---DQVVLSYGGL-RGAVA 419

Qy     397 IALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHS 456
           ||          : |  : :::|: || |: :  ||  |||:  |    |   | |
Db     420 FALV----VLLDGDKVKEKNLFVSTTIIVVFFTVIFQGLTIKPLVQWL----KVKRSEHR 471

Qy     457 SMTTSEPSSPKHFTVPLLDNQPDSESDMITGPEVARPTALRMLLRTPTHTVHRY----WR 512
              :|    : |   |      |  : |:|                    | |    |
Db     472 EPRLNEKLHGRAFDHIL------SAIEDISG-----------------QIGHNYLRDKWS 508

Qy     513 KFDDSFMRPV 522
           ||  |:  |
Db     509 HFDRKFLSRV 518


RESULT 14
NAH1_PIG
ID   NAH1_PIG       STANDARD;      PRT;   818 AA.
AC   P48762;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Sodium/hydrogen exchanger 1 (Na(+)/H(+) exchanger 1) (NHE-1).
GN   SLC9A1 OR NHE1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92087905; PubMed=1661081;
RA   Reilly R.F., Hildebrandt F., Biemesderfer D., Sardet C.,
RA   Pouyssegur J., Aronson P.S., Slayman C.W., Igarashi P.;
RT   "cDNA cloning and immunolocalization of a Na(+)-H+ exchanger in
RT   LLC-PK1 renal epithelial cells.";
RL   Am. J. Physiol. 261:F1088-F1094(1991).
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC       TRANSDUCTION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC       DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC   -!- CAUTION: HYDROPHOBIC DOMAINS A, B AND L ARE NOT BELIEVED TO BE
CC       TRANSMEMBRANAL, BUT ONLY MEMBRANE-ASSOCIATED.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M89631; AAA31092.1; -.
DR   EMBL; S71135; AAB20633.1; -.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation.
FT   DOMAIN        1     11       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       12     31       A (M1) HYDROPHOBIC.
FT   DOMAIN       32    101       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      102    123       B (M2) HYDROPHOBIC.
FT   DOMAIN      124    126       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    127    146       C (M3) (POTENTIAL).
FT   DOMAIN      147    158       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    159    179       D (M4) (POTENTIAL).
FT   DOMAIN      180    184       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    185    206       E (M5) (POTENTIAL).
FT   DOMAIN      207    226       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    227    247       F (M5A) (POTENTIAL).
FT   DOMAIN      248    256       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    257    278       G (M5B) (POTENTIAL).
FT   DOMAIN      279    297       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    298    318       H (M6) (POTENTIAL).
FT   DOMAIN      319    332       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    333    353       I (M7) (POTENTIAL).
FT   DOMAIN      354    384       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    385    406       J (M8) (POTENTIAL).
FT   DOMAIN      407    412       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    413    434       K (M9) (POTENTIAL).
FT   DOMAIN      435    448       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      449    469       L, HYDROPHOBIC.
FT   DOMAIN      470    478       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    479    499       M (M10) (POTENTIAL).
FT   DOMAIN      500    818       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    370    370       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    683    683       H -> Y (IN REF. 1; AAB20633).
SQ   SEQUENCE   818 AA;  90987 MW;  9329F7D9A51D3DC9 CRC64;

  Query Match             16.6%;  Score 460.5;  DB 1;  Length 818;
  Best Local Similarity   28.0%;  Pred. No. 3e-22;
  Matches  145;  Conservative  108;  Mismatches  181;  Indels   83;  Gaps   22;

Qy      25 MNLFVALLCACIVLGHLLEE-NRWVNESITALIIGLCTGVVILLLSGGKSSHLLVFSEDL 83
           ::|:: | |   :  |::   :  | ||   :::||  | :|  :  |::   |    ::
Db     105 ISLWILLACLMKIGFHVIPTISSIVPESCLLIVVGLLVGGLIKAV--GETPPFL--QSEV 160

Qy      84 FFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSII--FGAVKIFKHLDID 141
           ||::|||||| :||: :  :||  |  ||::|  :||| :   :     || :     |:
Db     161 FFLFLLPPIILDAGYFLPLRQFTENLGTILIFAVVGTLWNAFFLGGLMYAVCLVGGEQIN 220

Qy     142 FLDFGDYLAIGAIFAATDSVCTLQVLSQDE-TPLLYSLVFGEGVVNDATSVVLFNAIQSF 200
           :   | |  |:| :| | |  | |  :     ||: ||||| ::||| :|||::  : |
Db     221 NIGLLDNLLFGSIISAVDPVAVLAVFEEIHINELLHILVFGESLLNDAVTVVLYHLFEEF 280
```

```
Qy     201 DMTSFDPKIG-----LHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDREVAL 255
           ::| ::|     | |: :|  : |   |:||  |:: |:  :   |  |     |
Db     281 --ANYD-RVGIVDIVLGFL-SFFVVSLGGVFVGVVYGVIAAFTSR---FTSHIRVIEPLF 333

Qy     256 MMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETF 315
           : | ||::|: ||||:|||| :   |:||  |   |::  |  | ::     | |:||
Db     334 VFLYSYMAYLSAELFHLSGIMALIASGVVMRPYVEANISHKSHTTIKYFLKMWSSVSETL 393

Qy     316 IFLYVGMDALDIE---KWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKKN 372
           ||:::|:  :       | ||         :|::|    |: |   |  |::  |   |
Db     394 IFIFLGVSTVAGSHHWNWTFV---------ISTLL--FCLIARVLGVLGLTWFIN---KF 439

Qy     373 SSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHEN-----AIMITSTVTVVL 427
           |:: : | || : || |||:: :|        ||  | :|      : :|: :||:
Db     440 RIVKLTPKDQFIIAYGGL-RGAIAFSL--------GHL-LDKNHFPMCDLFLTAIITVIF 489

Qy     428 FSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMITG 487
           |:  | |:  :||::||    ||          :: |  :      ||:       ::||
Db     490 FTVFVQGMTIRPLVDLLAVKKKQ---------ETKRSINEEIHTQFLDH-------LLTG 533

Qy     488 PEVARPTALRMLLRTPTHTVHRYWR----KFDDSFMR 520
           |              |  | :|:    :|:  :::
Db     534 IE-----------DICGHYGHHHWKDKLNRFNKKYVK 559


RESULT 15
NAH1_RABIT
ID   NAH1_RABIT     STANDARD;      PRT;   816 AA.
AC   P23791;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Sodium/hydrogen exchanger 1 (Na(+)/H(+) exchanger 1) (NHE-1).
GN   SLC9A1 OR NHE1.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NEW ZEALAND WHITE; TISSUE=Ileal villus;
RX   MEDLINE=91293066; PubMed=1712287;
RA   Tse C.-M., Ma A.I., Yang V.W., Watson A.J.M., Levine S.,
RA   Montrose M.H., Potter J., Sardet C., Pouyssegur J., Donowitz M.;
RT   "Molecular cloning and expression of a cDNA encoding the rabbit ileal
RT   villus cell basolateral membrane Na+/H+ exchanger.";
RL   EMBO J. 10:1957-1967(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=92096447; PubMed=1661611;
RA   Hildebrandt F., Pizzonia J.H., Reilly R.F., Reboucas N.A.,
RA   Sardet C., Pouyssegur J., Slayman C.W., Aronson P.S., Igarashi P.;
RT   "Cloning, sequence, and tissue distribution of a rabbit renal Na+/H+
RT   exchanger transcript.";
RL   Biochim. Biophys. Acta 1129:105-108(1991).
RN   [3]
RP   SEQUENCE OF 472-816 FROM N.A.
RC   STRAIN=NEW ZEALAND WHITE; TISSUE=Heart muscle;
RX   MEDLINE=91138752; PubMed=1704856;
RA   Fliegel L., Sardet C., Pouyssegur J., Barr A.;
RT   "Identification of the protein and cDNA of the cardiac Na+/H+
RT   exchanger.";
RL   FEBS Lett. 279:25-29(1991).
CC   -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC       BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC       CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC       SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC       TRANSDUCTION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: KIDNEY AND INTESTINE.
CC   -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC   -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC   -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC       DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X59935; CAA42558.1; -.
DR   EMBL; X61504; CAA43721.1; -.
DR   EMBL; X56536; CAA39881.1; -.
DR   PIR; S13926; S13926.
DR   PIR; S16328; S16328.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
KW   Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW   Multigene family; Phosphorylation.
FT   DOMAIN        1     15       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     16     35       M1 (POTENTIAL).
FT   DOMAIN       36    107       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    108    127       M2 (POTENTIAL).
FT   DOMAIN      128    129       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    130    149       M3 (POTENTIAL).
FT   DOMAIN      150    154       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    155    174       M4 (POTENTIAL).
FT   DOMAIN      175    191       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    192    211       M5 (POTENTIAL).
FT   DOMAIN      212    227       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    228    247       M5A (POTENTIAL).
FT   DOMAIN      248    256       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    257    276       M5B (POTENTIAL).
FT   DOMAIN      277    294       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    295    315       M6 (POTENTIAL).
FT   DOMAIN      316    338       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    339    358       M7 (POTENTIAL).
FT   DOMAIN      359    386       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    387    406       M8 (POTENTIAL).
FT   DOMAIN      407    410       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    411    430       M9 (POTENTIAL).
FT   DOMAIN      431    480       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    481    500       M10 (POTENTIAL).
FT   DOMAIN      501    816       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     75     75       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    370    370       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    242    242       V -> A (IN REF. 2).
FT   CONFLICT    569    569       K -> E (IN REF. 2).
SQ   SEQUENCE   816 AA;  90717 MW;  336738D267F7F436 CRC64;

  Query Match             16.6%;  Score 459.5;  DB 1;  Length 816;
  Best Local Similarity   28.2%;  Pred. No. 3.5e-22;
  Matches  144;  Conservative  105;  Mismatches  191;  Indels   71;  Gaps   20;

Qy      25 MNLFVALLCACIVLGHLLEE-NRWVNESITALIIGLCTGVVILLLSGGKSSHLLVFSEDL 83
           ::|:: | |   :  |::   :  | ||   :::||  | :|  : | |   |     ::
Db     105 ISLWILLACLMKIGFHVIPTISSIVPESCLLIVVGLLVGGLIKGV-GEKPPFL---QSEV 160

Qy      84 FFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIIS--FGAVKIFKHLDID 141
           ||::|||||| :||: :  :||  |  ||::|  :||| :   :     || :     |:
Db     161 FFLFLLPPIILDAGYFLPLRQFTENLGTILIFAVVGTLWNAFFLGGLMYAVCLVGGEQIN 220

Qy     142 FLDFGDYLAIGAIFAATDSVCTLQVLSQDE-TPLLYSLVFGEGVVNDATSVVLFNAIQSF 200
             :  | |  |:| :| | |  | |  :     ||: ||||| ::||| :|||::  : |
Db     221 NIGLLDNLLFGSIISAVDPVAVLAVFEEIHINELLHILVFGESLLNDAVTVVLYHLFEEF 280

Qy     201 DMTSFD----PKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDREVALM 256
             ::|      | | |: :|  : |   |:||  |:: |:  :   |  |    |  :
Db     281 --ANYDHVGIVDIVLGFL-SFFVVALGGVFVGVVYGVIAAFTSR---FTAHIRVIEPLFV 334
```

```
Qy     257 MLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETFI 316
           | ||::|: ||||:||||: :   |:||  |   |::  |  | ::     | |:|| |
Db     335 FLYSYMAYLSAELFHLSGIMALIASGVVMRPYVEANISHKSHTTIKYFLKMWSSVSETLI 394

Qy     317 FLYVGMDALDIE---KWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKKNS 373
           |:::|:  :       | ||         :|::|    |: |   |  |::  |   |
Db     395 FIFLGVSTVAGSHHWNWTFV---------ISTLL--FCLIARVLGVLGLTWFIN---KFR 440

Qy     374 SDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTVVF 433
             |:: : | || : || |||:: :| |       |  : :  : :|: :||: |:  |
Db     441 IVKLTPKDQFIIAYGGL-RGAIAFSLGY--LLDKKHFPMCD--LFLTAIITVIFFTVFVQ 495

Qy     434 GLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMITGPEVARP 493
           |:  :||::||    ||          :: |  :      ||:       ::|| |
Db     496 GMTIRPLVDLLAVKKKQ---------ETKRSINEEIHTQFLDH-------LLTGIE---- 535

Qy     494 TALRMLLRTPTHTVHRYWR----KFDDSFMR 520
                 |  | :|:    :|:  :::
Db     536 -------DICGHYGHHHWKDKLNRFNKKYVK 559
```

Search completed: October 18, 2002, 12:26:10
Job time : 19 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: October 18, 2002, 12:24:10 ; Search time 15 Seconds
(without alignments)
882.578 Million cell updates/sec

Title: US-09-830-123-2
Perfect score: 2768
Sequence: 1 MAFGLSSLLQNSDLFTSDHA.....FGGRGFVPFVAGSPVEQSPR 542

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	471.5	17.0	832	2	US-08-677-734A-12	Sequence 12, Appl
2	468.5	16.9	831	2	US-08-677-734A-11	Sequence 11, Appl
3	461	16.7	834	2	US-08-677-734A-9	Sequence 9, Appli
4	461	16.7	834	2	US-08-677-734A-10	Sequence 10, Appl
5	112.5	4.1	438	2	US-08-677-049-9	Sequence 9, Appli
6	109.5	4.0	1394	4	US-09-213-053-2	Sequence 2, Appli
7	106	3.8	635	2	US-09-014-969-11	Sequence 11, Appl
8	100	3.6	775	2	US-08-714-070A-1	Sequence 1, Appli
9	99	3.6	584	2	US-08-928-692-13	Sequence 13, Appl
10	97.5	3.5	352	3	US-08-466-343D-2	Sequence 2, Appli
11	97.5	3.5	352	4	US-09-087-232A-13	Sequence 13, Appl
12	97.5	3.5	352	4	US-08-861-105-14	Sequence 14, Appl
13	97.5	3.5	352	4	US-08-575-967A-2	Sequence 2, Appli
14	97.5	3.5	352	4	US-09-045-583-52	Sequence 52, Appl
15	96	3.5	664	4	US-09-268-140-2	Sequence 2, Appli
16	95.5	3.5	370	4	US-09-144-914-8	Sequence 8, Appli
17	95	3.4	1786	2	US-08-477-451-16	Sequence 16, Appl
18	93.5	3.4	397	1	US-08-098-141-2	Sequence 2, Appli
19	92.5	3.3	326	4	US-08-986-768-2	Sequence 2, Appli
20	92.5	3.3	513	4	US-09-097-889-15	Sequence 15, Appl
21	92.5	3.3	3287	2	US-08-477-451-7	Sequence 7, Appli
22	91.5	3.3	299	2	US-09-097-759-3	Sequence 3, Appli
23	91.5	3.3	306	2	US-09-097-759-2	Sequence 2, Appli
24	91.5	3.3	649	4	US-08-800-291B-6	Sequence 6, Appli
25	91.5	3.3	650	4	US-08-800-291B-4	Sequence 4, Appli
26	91	3.3	342	3	US-08-852-824-2	Sequence 2, Appli
27	90.5	3.3	360	4	US-08-875-573-20	Sequence 20, Appl
28	90.5	3.3	360	4	US-09-232-878-2	Sequence 2, Appli
29	90.5	3.3	360	4	US-09-045-583-55	Sequence 55, Appl
30	90.5	3.3	461	2	US-08-672-814D-2	Sequence 2, Appli
31	90.5	3.3	461	4	US-09-333-696-2	Sequence 2, Appli
32	90	3.3	266	2	US-08-225-480-7	Sequence 7, Appli
33	90	3.3	266	2	US-09-097-759-4	Sequence 4, Appli
34	90	3.3	916	2	US-08-928-692-58	Sequence 58, Appl
35	89.5	3.2	694	2	US-08-895-522-4	Sequence 4, Appli
36	89.5	3.2	694	3	US-09-195-391-4	Sequence 4, Appli
37	89	3.2	355	1	US-08-461-244-2	Sequence 2, Appli
38	89	3.2	355	4	US-09-045-583-56	Sequence 56, Appl
39	88	3.2	506	4	US-09-066-047-7	Sequence 7, Appli
40	88	3.2	524	2	US-08-928-692-12	Sequence 12, Appl
41	88	3.2	1334	2	US-08-996-545-2	Sequence 2, Appli
42	88	3.2	1334	4	US-09-328-320-2	Sequence 2, Appli
43	87.5	3.2	411	4	US-09-236-080-6	Sequence 6, Appli
44	87.5	3.2	649	4	US-08-800-291B-5	Sequence 5, Appli
45	87	3.1	459	4	US-09-097-889-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-08-677-734A-12
; Sequence 12, Application US/08677734A
; Patent No. 5871919
; GENERAL INFORMATION:
; APPLICANT: Brant, Steven R.
; APPLICANT: Yun, Chris C.H.
; APPLICANT: Donowitz, Mark
; APPLICANT: Tse, Chung-Ming
; TITLE OF INVENTION: Cloning, Tissue Distribution, and
; TITLE OF INVENTION: Functional Analysis Of The Human Na+/H+ Exchanger Isoform,
; TITLE OF INVENTION: NHE3.
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/677,734A
; FILING DATE: 10-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 05387.0043-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 832 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-677-734A-12

Query Match 17.0%; Score 471.5; DB 2; Length 832;
Best Local Similarity 28.4%; Pred. No. 3.6e-38;
Matches 147; Conservative 92; Mismatches 200; Indels 79; Gaps 18;

```
Qy      22 VVSMNLFVALLCACIV-LGHLLEENRWVNESITALIIGLCTGVVILLLSGGKSSHLLVF- 79
           :::: : || |   :  | |  :    | ||   :::||  | ::|      : |:  |
Db      56 IIALWVLVASLAKIVFHLSH--KVTSVVPESALLIVLGLVLGGIVL-----AADHIASFT 108

Qy      80 -SEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIISFGAVKIF--- 135
            :  :|| ||||||: :||: :  : || |  :|:|:  :||: : :        :|
Db     109 LTPTVFFFYLLPPIVLDAGYFMPNRLFFSNLGSILLYAVVGTVWNAATTGLSLYGVFLSG 168

Qy     136 --KHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQ-DETPLLYSLVFGEGVVNDATSVV 192
             | |  |||   |  |:: || | |  | |  :     :|: :|||| ::||| :||
Db     169 IMGELKIGLLDF---LLFGSLIAAVDPVAVLAVFEEVHVNEVLFIIVFGESLLNDAVTVV 225

Qy     193 LFNAIQSFDMTSFDPKIGLHF---IGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHST 249
           |:|  |||     |   |:     | :|  : |  | :||    | : : :   | :|
Db     226 LYNVFQSFVTLGGDKVTGVDCVKGIVSFFVVSLGGTLVGVVFAFLLSLVTR---FTKHVR 282

Qy     250 DREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLS 309
             |   : ::|||||: :|:  || || : ||||    |   |::| |  | |::   |:
Db     283 VIEPGFVFIISYLSYLTSEMLSLSSILAITFCGICCQKYVKANISEQSATTVRYTMKMLA 342

Qy     310 FVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLA 369
             ||| ||:::|: |:|   | :      : |   : :  : | ||  |   ::| |
Db     343 SGAETIIFMFLGISAVDPLIWTW-------NTAFVRLTLLFVSVFRAIGVVLQTWLLNRY 395

Qy     370 KKNSSDKISFRQQIIIWWAGLMRGAVSIALAY----NKFTTSGHTSLHENAIMITSTVTV 425
           :   : |    |::: : || ||||: ||      ||        : |  : :::|: |
Db     396 RMVQLELI---DQVVMSYGGL-RGAVAFALVALLDGNK--------VKEKNLFVSTTIIV 443

Qy     426 VLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMI 485
           | |: :  ||  |||:  |    |   | |     :|    : | `|     |  : |
Db     444 VFFTVIFQGLTIKPLVQWL----KVKRSEHREPKLNEKLHGRAFDHIL------SAIEDI 493

Qy     486 TGPEVARPTALRMLLRTPTHTVHRY----WRKFDDSFM 519
           :|                  | |    |  ||  |:
Db     494 SG-----------------QIGHNYLRDKWANFDRRFL 514
```

```
RESULT 2
US-08-677-734A-11
; Sequence 11, Application US/08677734A
; Patent No. 5871919
;  GENERAL INFORMATION:
;    APPLICANT:  Brant, Steven R.
;    APPLICANT:  Yun, Chris C.H.
;    APPLICANT:  Donowitz, Mark
;    APPLICANT:  Tse, Chung-Ming
;    TITLE OF INVENTION:  Cloning, Tissue Distribution, and
;    TITLE OF INVENTION:  Functional Analysis Of The Human Na+/H+ Exchanger Isoform,
;    TITLE OF INVENTION:  NHE3.
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W., Suite 700
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/677,734A
;      FILING DATE:  10-JUL-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fordis, Jean B.
;      REGISTRATION NUMBER:  32,984
;      REFERENCE/DOCKET NUMBER:  05387.0043-00000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 408-4000
;      TELEFAX:  (202) 408-4400
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  831 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-677-734A-11

  Query Match             16.9%;  Score 468.5;  DB 2;  Length 831;
  Best Local Similarity   27.7%;  Pred. No. 7.2e-38;
  Matches  146;  Conservative   89;  Mismatches  200;  Indels   93;  Gaps   18;

Qy      22 VVSMNLFVALLCACIV-LGHLLEENRWVNESITALIIGLCTGVVILLLSGGKSSHLLVF- 79
           :::: : || |   :  | |  :    | ||   :::||  | ::|      : |:  |
Db      54 IIALWILVASLAKIVFHLSH--KVTSVVPESALLIVLGLVLGGIV-----WAADHIASFT 106

Qy      80 -SEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIISFGAVKIF--- 135
            :  ||| ||||||: :||: :  : || |  ||:|:  |||: : :        :|
Db     107 LTPTLFFFYLLPPIVLDAGYFMPNRLFFGNLGTILLYAVIGTIWNAATTGLSLYGVFLSG 166

Qy     136 --KHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQ-DETPLLYSLVFGEGVVNDATSVV 192
              | |  |||   |  |:: || | |  | |  :     :|: :|||| ::||| :||
Db     167 LMGELKIGLLDF---LLFGSLIAAVDPVAVLAVFEEVHVNEVLFIIVFGESLLNDAVTVV 223

Qy     193 LFNAIQSFDMTSFDPKIGLHF---IGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHST 249
           |:|  :||     |   |:     | :|  : |  | :||    | : : :   | :|
Db     224 LYNVFESFVTLGGDAVTGVDCVKGIVSFFVVSLGGTLVGVIFAFLLSLVTR---FTKHVR 280

Qy     250 DREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLS 309
             |   : ::|||||: :|:  || || : ||||    |   |::| |  | |::   |:
Db     281 IIEPGFVFVISYLSYLTSEMLSLSAILAITFCGICCQKYVKANISEQSATTVRYTMKMLA 340

Qy     310 FVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLA 369
             ||| ||:::|: |:|   | :      : |   : :  | | ||  |   ::: |
Db     341 SGAETIIFMFLGISAVDPVIWTW-------NTAFVLLTLVFISVYRAIGVVLQTWILNRY 393

Qy     370 KKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFS 429
           :   : |    |::: : || ||||: ||           : |  : :::|: || |:
Db     394 RMVQLETI---DQVVMSYGGL-RGAVAYALV----VLLDEKKVKEKNLFVSTTLIVVFFT 445

Qy     430 TVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPK----------HFTVPLLDNQP 478
           :  ||  |||:  |             :  ||  ||              |    : |
Db     446 VIFQGLTIKPLVQWL-----------KVKRSEQREPKLNEKLHGRAFDHILSAIED--- 490

Qy     479 DSESDMITGPEVARPTALRMLLRTPTHTVHRY----WRKFDDSFMRPV 522
                 |:|                  | |    |  ||  |:  |
Db     491 ------ISG----------------QIGHNYLRDKWSNFDRKFLSKV 515
```

```
RESULT 3
US-08-677-734A-9
; Sequence 9, Application US/08677734A
; Patent No. 5871919
;  GENERAL INFORMATION:
;    APPLICANT:  Brant, Steven R.
;    APPLICANT:  Yun, Chris C.H.
;    APPLICANT:  Donowitz, Mark
;    APPLICANT:  Tse, Chung-Ming
;    TITLE OF INVENTION:  Cloning, Tissue Distribution, and
;    TITLE OF INVENTION:  Functional Analysis Of The Human Na+/H+ Exchanger Isoform,
;    TITLE OF INVENTION:  NHE3.
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W., Suite 700
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/677,734A
;       FILING DATE:  10-JUL-1996
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Fordis, Jean B.
;       REGISTRATION NUMBER:  32,984
;       REFERENCE/DOCKET NUMBER:  05387.0043-00000
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 408-4000
;       TELEFAX:  (202) 408-4400
;   INFORMATION FOR SEQ ID NO:  9:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  834 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-677-734A-9

  Query Match             16.7%;  Score 461;  DB 2;  Length 834;
  Best Local Similarity   28.6%;  Pred. No. 4e-37;
  Matches  140;  Conservative   85;  Mismatches  197;  Indels   68;  Gaps   15;

Qy      48 VNESITALIIGLCTGVVILLLSGGKSSHLLVF--SEDLFFIYLLPPIIFNAGFQVKKKQF 105
           | ||   :::||  | ::       : |:  |  :  :|| ||||||: :||: :  : |
Db      82 VPESALLIVLGLVLGGIV-----WAADHIASFTLTPTVFFFYLLPPIVLDAGYFMPNRLF 136

Qy     106 FVNFMTIMLFGAIGTLISCSIISFGAVKIF-----KHLDIDFLDFGDYLAIGAIFAATDS 160
           | |  ||:|:  :||: : :          :|       | |  |||   |  |:: || |
Db     137 FGNLGTILLYAVVGTVWNAATTGLSLYGVFLSGLMGDLQIGLLDF---LLFGSLMAAVDP 193

Qy     161 VCTLQVLSQ-DETPLLYSLVFGEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHF---IGN 216
           |  | |  :     :|: :|||| ::||| :|||:|  :||      |   |:     | :
Db     194 VAVLAVFEEVHVNEVLFIIVFGESLLNDAVTVVLYNVFESFVALGGDNVTGVDCVKGIVS 253

Qy     217 FLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGIL 276
           |  : |  | :||    | : : :   | :|    |   : ::|||||: :|:  || ||
Db     254 FFVVSLGGTLVGVVFAFLLSLVTR---FTKHVRIIEPGFVFIISYLSYLTSEMLSLSAIL 310

Qy     277 TVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNS 336
           : ||||    |   |::| |  | |::   |:  ||| ||:::|: |::   | :
Db     311 AITFCGICCQKYVKANISEQSATTVRYTMKMLASSAETIIFMFLGISAVNPFIWTW---- 366

Qy     337 QGLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVS 396
           : |   : :  | | ||  |   ::| |  :    : |    |::: : || ||||:
Db     367 ---NTAFVLLTLVFISVYRAIGVVLQTWLLNRYRMVQLEPI---DQVVLSYGGL-RGAVA 419

Qy     397 IALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHS 456
           ||           : |  : :::|: || |: :  ||  |||:  |    |   | |
Db     420 FALV----VLLDGDKVKEKNLFVSTTIIVVFFTVIFQGLTIKPLVQWL----KVKRSEHR 471

Qy     457 SMTTSEPSSPKHFTVPLLDNQPDSESDMITGPEVARPTALRMLLRTPTHTVHRY----WR 512
           :|    : |   |      |  : |:|                  | |    |
Db     472 EPRLNEKLHGRAFDHIL------SAIEDISG----------------QIGHNYLRDKWS 508

Qy     513 KFDDSFMRPV 522
           ||  |:  |
Db     509 HFDRKFLSRV 518


RESULT 4
US-08-677-734A-10
; Sequence 10, Application US/08677734A
; Patent No. 5871919
;  GENERAL INFORMATION:
;     APPLICANT:  Brant, Steven R.
;     APPLICANT:  Yun, Chris C.H.
;     APPLICANT:  Donowitz, Mark
;     APPLICANT:  Tse, Chung-Ming
;     TITLE OF INVENTION:  Cloning, Tissue Distribution, and
;     TITLE OF INVENTION:  Functional Analysis Of The Human Na+/H+ Exchanger Isoform,
;     TITLE OF INVENTION:  NHE3.
;     NUMBER OF SEQUENCES:  12
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;       ADDRESSEE:  Dunner
;       STREET:  1300 I Street, N.W., Suite 700
;       CITY:  Washington
;       STATE:  D.C.
;       COUNTRY:  USA
;       ZIP:  20005-3315
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/677,734A
;       FILING DATE:  10-JUL-1996
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Fordis, Jean B.
;       REGISTRATION NUMBER:  32,984
;       REFERENCE/DOCKET NUMBER:  05387.0043-00000
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 408-4000
;       TELEFAX:  (202) 408-4400
;   INFORMATION FOR SEQ ID NO:  10:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  834 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-677-734A-10

  Query Match             16.7%;  Score 461;  DB 2;  Length 834;
  Best Local Similarity   28.6%;  Pred. No. 4e-37;
  Matches  140;  Conservative   85;  Mismatches  197;  Indels   68;  Gaps   15;

Qy      48 VNESITALIIGLCTGVVILLLSGGKSSHLLVF--SEDLFFIYLLPPIIFNAGFQVKKKQF 105
           | ||   :::||  | ::       : |:  |  :  :|| ||||||: :||: :  : |
Db      82 VPESALLIVLGLVLGGIV-----WAADHIASFTLTPTVFFFYLLPPIVLDAGYFMPNRLF 136

Qy     106 FVNFMTIMLFGAIGTLISCSIISFGAVKIF-----KHLDIDFLDFGDYLAIGAIFAATDS 160
           | |  ||:|:  :||: : :          :|       | |  |||   |  |:: || |
Db     137 FGNLGTILLYAVVGTVWNAATTGLSLYGVFLSGLMGDLQIGLLDF---LLFGSLMAAVDP 193

Qy     161 VCTLQVLSQ-DETPLLYSLVFGEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHF---IGN 216
           |  | |  :     :|: :|||| ::||| :|||:|  :||      |   |:     | :
Db     194 VAVLAVFEEVHVNEVLFIIVFGESLLNDAVTVVLYNVFESFVALGGDNVTGVDCVKGIVS 253

Qy     217 FLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGIL 276
           |  : |  | :||    | : : :   | :|    |   : ::|||||: :|:  || ||
Db     254 FFVVSLGGTLVGVVFAFLLSLVTR---FTKHVRIIEPGFVFIISYLSYLTSEMLSLSAIL 310

Qy     277 TVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNS 336
           : ||||    |   |::| |  | |::   |:  ||| ||:::|: |::   | :
Db     311 AITFCGICCQKYVKANISEQSATTVRYTMKMLASSAETIIFMFLGISAVNPFIWTW---- 366

Qy     337 QGLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVS 396
           : |   : :  | | ||  |   ::| |  :    : |    |::: : || ||||:
Db     367 ---NTAFVLLTLVFISVYRAIGVVLQTWLLNRYRMVQLEPI---DQVVLSYGGL-RGAVA 419

Qy     397 IALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHS 456
           ||           : |  : :::|: || |: :  ||  |||:  |    |   | |
Db     420 FALV----VLLDGDKVKEKNLFVSTTIIVVFFTVIFQGLTIKPLVQWL----KVKRSEHR 471

Qy     457 SMTTSEPSSPKHFTVPLLDNQPDSESDMITGPEVARPTALRMLLRTPTHTVHRY----WR 512
           :|    : |   |      |  : |:|                  | |    |
Db     472 EPRLNEKLHGRAFDHIL------SAIEDISG----------------QIGHNYLRDKWS 508
```

```
Qy     513 KFDDSFMRPV 522
           ||   |:  |
Db     509 HFDRKFLSRV 518


RESULT 5
US-08-677-049-9
; Sequence 9, Application US/08677049
; Patent No. 5858707
;  GENERAL INFORMATION:
;    APPLICANT:  Guimaraes, M. Jorge
;    APPLICANT:  Bazan, J. Fernando
;    APPLICANT:  McClanahan, Terrill K.
;    APPLICANT:  Zlotnik, Albert
;    TITLE OF INVENTION:  PURIFIED MAMMALIAN NUCLEOBASE PERMEASES;
;    TITLE OF INVENTION:  NUCLEIC ACIDS; ANTIBODIES
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DNAX Research Institute
;      STREET:  901 California Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/677,049
;      FILING DATE:  03-JUL-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/000,788
;      FILING DATE:  03-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ching, Edwin P.
;      REGISTRATION NUMBER:  34,090
;      REFERENCE/DOCKET NUMBER:  DX0511
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-852-9196
;      TELEFAX:  415-496-1200
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  438 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  115..144
;      OTHER INFORMATION:  /note= "Encompasses TM 4 of Figure
;      OTHER INFORMATION:  4"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  323..357
;      OTHER INFORMATION:  /note= "Encompasses TM 9 of Figure
;      OTHER INFORMATION:  4"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  359..386
;      OTHER INFORMATION:  /note= "Encompasses TM 10 of Figure
;      OTHER INFORMATION:  4"
US-08-677-049-9

  Query Match             4.1%;  Score 112.5;  DB 2;  Length 438;
  Best Local Similarity   22.1%;  Pred. No. 0.0076;
  Matches  103;  Conservative   69;  Mismatches  153;  Indels  141;  Gaps   25;

Qy      56 IIGLCTGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAG----FQVKKKQFFVNFMT 111
           :: :  | ::: |  ||:  | |  | |   ||:   ||  |     ||   :||   :
Db      16 VLAMYAGAIVVPLIVGKAMGLTV--EQL--TYLVSIDIFMCGVATLLQVWSNRFFGIGLP 71

Qy     112 IML---FGAIGTLISC----------SIISFGAVKIFKHLDIDFLDFGDYLAIGAIF-- 155
           ::|   | |:  :|:           |||: | : |     | |  ||  :   : |
Db      72 VVLGCTFTAVSPMIAIGSEYGVSTVYGSIIASGILVIL----ISFF-FGKLV---SFFPP 123

Qy     156 AATDSVCTLQVLSQDETPL-LYSLVFGEGVVN--DATSVVLFNAIQSFDMTSFDPKIGLH 212
             | || |  ::     |: : ::  |||  :  | ::: |   : |  :
Db     124 VVTGSVVT--IIGITLMPVAMNNMAGGEGSADFGDLSNLALAFTVLSIIV---------- 171

Qy     213 FIGNFLYLFLS------STFLGVGIGLLCAYIIKKLYFGRHSTDREVALMMLMSYLSYIM 266
             || |        |  :|: ||   || : |: |   |     | :: |   |
Db     172 ----LLYRFTKGFIKSVSILIGILIGTFIAYFMGKVQFDNVSD----AAVVQMIQPFYFG 223

Qy     267 AELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETFIFLYVG----- 321
           |  |: : |:|:    ||                          | |  | ::  :|
Db     224 APSFHAAPIITMSIVAIV----------------------SLVESTGVYFALGDLTNR 259

Qy     322 -MDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFP-LSFLSNLA-------KKN 372
            :  :|: |      ::||:|          |:|     ||  :|  |:        |||
Db     260 RLTEIDLSKG---YRAEGLAV----------LLGGIFNAFPYTAFSQNVGLVQLTGIKKN 306

Qy     373 S-------------SDKISFRQQII--IWWAGLMRGAVSIALAYN-KFTTSGHTSLHEN 415
           :                ||:   ||     | |     : :||  |  :    :  ||
Db     307 AVIVVTGVILMAFGLFPKIAAFTTIIPSAVLGGAMVAMFGMVIAYGIKMLSRIDFAKQEN 366

Qy     416 AIMITSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTS 461
            :::  :| : |  |||  :             ||:||  : :||:
Db     367 LLIVACSVGLGLGVTVVPDIF-----------KQLPSALTLLTTN 400


RESULT 6
US-09-213-053-2
; Sequence 2, Application US/09213053
; Patent No. 6159477
; GENERAL INFORMATION:
;  APPLICANT: AUDONNET, Jean-Christophe
;  APPLICANT: BAUDU, Philippe
;  TITLE OF INVENTION: CANINE HERPESVIRUS BASED RECOMBINANT LIVE VACCINE, IN
;  TITLE OF INVENTION: PARTICULAR AGAINST CANINE DISTEMPER, RABIES OR THE
;  TITLE OF INVENTION: PARAINFLUENZA 2 VIRUS
;  FILE REFERENCE: 454313-2200
;  CURRENT APPLICATION NUMBER: US/09/213,053
;  CURRENT FILING DATE: 1998-12-16
;  EARLIER APPLICATION NUMBER: 9608242
;  EARLIER FILING DATE: 1996-06-27
;  EARLIER APPLICATION NUMBER: PCT/FR97/01115
;  EARLIER FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 1394
;   TYPE: PRT
;   ORGANISM: Canine herpesvirus
US-09-213-053-2

  Query Match             4.0%;  Score 109.5;  DB 4;  Length 1394;
  Best Local Similarity   20.4%;  Pred. No. 0.087;
  Matches  114;  Conservative   84;  Mismatches  195;  Indels  167;  Gaps   26;

Qy      21 SVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLCTGVVILLLS------GGKSS 74
           |:: : |   :| | |:|  :|       ||    || : | |:|  ||       | |
Db     463 SLLCIGLQAGILTALIILIQIL------TESFVCSIILIAT-VLIFTLSKISISTSEKIS 515

Qy      75 HLLVFSEDLF---------FIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCS 125
             :   |: :|         | ::| ||
Db     516 SICRISQSIFVTIAAFCWGFDWILNPI-------------------------------- 542

Qy     126 IISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVFGEGV- 184
                |:||   | : ||               ::||:::       | |:: | |
Db     543 -----AIKIILILSLSFL---------------TICTIKI----HIFYLISILNGSGSH 577
```

```
Qy    185 VNDATSVVLFNAI----QSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIK 240
Db    578 VKGSLVTILFGTILGVFGTLNVIKIEILIG---FGIALCIILSNTNFGLVIRDTCYYRIG 634

Qy    241 KLYFGRHSTD----------------------REV-ALMMLMSYLSYIMAELFYLSGILT 277
Db    635 RYKLMRTFTDLGHGASYSIEEDETSDYSEIHERKISSFQLIYKYPSMIIISI--LGFMLT 692

Qy    278 VFFCG--IVMSHYTWHNVTESSRVTTRHSFAT---LSFVAETFIFLYVGMDALDIEKWKF 332
Db    693 IAIWGLNVYLKNLKFH-----SPFTLVISFIVGHCLAFLVEPFNY----------KIKC 736

Qy    333 VKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMR 392
Db    737 ISRIILIICLLLELIASLISVLGLNFGSPL-ILTTTTTISLVSLLYIRKQT----QGVNR 791

Qy    393 GA---VSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTVVFGLMTKPLINLLLPPHK 449
Db    792 LAATYISRALIIGLYMTVGICYIFIKTINMSFKNFYLIYVIIIF-------INSII---- 840

Qy    450 QMPSGHSSMTTSEPSSPKHFTVPLLDNQ-PDSESDMITGPEVARPTALRML---LRTPTH 505
Db    841 -----TSASTSKPSTPT--IIPTSANESPASIDTTITKPISTEANNLKSVSTSIKPPKN 892

Qy    506 TVHRYWRK--FDDSFMRPVF 523
Db    893 LKKKLLKSKCRDNVIYRPYF 912


RESULT 7
US-09-014-969-11
; Sequence 11, Application US/09014969
; Patent No. 5965397
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  McCoy, John M.
;    APPLICANT:  LaVallie, Edward R.
;    APPLICANT:  Racie, Lisa A.
;    APPLICANT:  Merberg, David
;    APPLICANT:  Treacy, Maurice
;    APPLICANT:  Spaulding, Vikki
;    APPLICANT:  Agostino, Michael J.
;    TITLE OF INVENTION:  SECRETED PROTEINS AND POLYNUCLEOTIDES
;    TITLE OF INVENTION:  ENCODING THEM
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/014,969
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sprunger, Suzanne A.
;      REGISTRATION NUMBER:  41,323
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8284
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  635 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-014-969-11

  Query Match             3.8%;  Score 106;  DB 2;  Length 635;
  Best Local Similarity   19.3%;  Pred. No. 0.059;
  Matches  106;  Conservative   83;  Mismatches  185;  Indels  176;  Gaps   25;

Qy     21 SVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIG----LCTGVVILLLSGGKSSHL 76
Db     25 SIMDYVVFVLLLVLSLAIGLYHACRGWGRHTVGELLMADRKMGCLPVALSLLATFQSA-- 82

Qy     77 LVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIISFGAVKIFK 136
Db     83 -------VAILRVPSEIYRFGTQYW-------FLRCCYF--LGLLIPAHIF----IPVFY 122

Qy    137 HLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVFGEGVVNDATSVVLFNA 196
Db    123 RLHL--TSAYEYLELR--FNKTVRVC-------GTVTFIFQMVIYMGVVLYAPSLAL-NA 170

Qy    197 IQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGV------GIGLLCAYIIKKLYFGRHSTD 250
Db    171 VTGFD-----------------LWLSVLALRIVCTVYTALGGLKAVI---------WTD 203

Qy    251 REVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVT---------T 301
Db    204 VFQTLVMFLGQLAVIIVGSAKVGGLGRV----------WAVASQHGRISGFELDPDPFV 252

Qy    302 RHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAV--------SSILVGLILV 353
Db    253 RHTFWTLAFGGVFMMLSLYGVNQAQVQRYLSSRTEKAAVLSCYAVFPFQQVSLCVG-CLI 311

Qy    354 GRAAFVFPLSFLSNLAKKNSSDKISFRQQIIIWW----------------AGLMRGAVS- 396
Db    312 GLVMFAYYQEYPMSIQQAQAAP-----DQFVLYFVMDLLKGLPGLPGLFIACLFSGSLST 366

Qy    397 IALAYNKFTTSGHTSL---------HENAIM-----------------ITSTVTVVLFS 429
Db    367 ISSAFNSLATVTMEDLIRPWFPEFSEARAIMLSRGLAFGYGLLCLGMAYISSQMGPVLQA 426

Qy    430 TV-VFGLMTKPLINL----LLPPHKQMPS--------------GHSSMTTS-------E 462
Db    427 AISIFGMVGGPLLGLFCLGMFFPCANPPGAVVGLLAGLVMAFWIGIGSIVTSMGFSMPPS 486

Qy    463 PSSPKHFTVP 472
Db    487 PSNGSSFSLP 496


RESULT 8
US-08-714-070A-1
; Sequence 1, Application US/08714070A
; Patent No. 5834237
;  GENERAL INFORMATION:
;    APPLICANT:  JACOBS, Eric
;    APPLICANT:  SILVESTRE, Nathalie
;    APPLICANT:  SCHWEINBRYBER, Ernst
;    TITLE OF INVENTION:  COMBINED USE OF TWO EXPRESSION CASSETTES
;    TITLE OF INVENTION:  FOR THE PRODUCTION OF A PROTEIN OF INTEREST
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  P.O. Box 1404
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/714,070A
```

```
;      FILING DATE:  07-OCT-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94/01767
;      FILING DATE:  10-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rea, Teresa Stanek
;      REGISTRATION NUMBER:  30,427
;      REFERENCE/DOCKET NUMBER:  017753-077
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  775 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Schizosaccharomyces pombe
US-08-714-070A-1

  Query Match             3.6%;  Score 100;  DB 2;  Length 775;
  Best Local Similarity   19.9%;  Pred. No. 0.32;
  Matches   74;  Conservative   65;  Mismatches  136;  Indels   96;  Gaps   16;

Qy    143 LDFGDYLAIG--------------AIFAATDSVCTLQVLSQDE------TPLLYSLVFGE 182
          : | | | :|                |       :   |: :|:|:      :| :    | |
Db    171 ISFRDQLKVGKLEDTLYLGPTTSEAFIERLQNELELESISEDDLYSKRLSPSVSYSEFDE 230

Qy    183 GVVNDATSVVLFNAIQSFDMTSFDPKIGL-------HFIGNFLYLFLSSTFLGVGIGLLC 235
           ::  | |::   |:  | : ||   :          ||    | :||:     :  | ||
Db    231 QLLLHARSLIPSKAVVEFLINSFFINVQTNLFVYHPHFFKCRLEIFLAME-NQIDAGFLC 289

Qy    236 AYIIKKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTE 295
            ::   :  ::: :::  :     : | |   ||  |  |::|   :: |     :  :
Db    290 ILLMVLAFGNQYTAEQQEDVSKSNFHASNIGNRLF--SAALSIFPLVLLQSDV---SAVQ 344

Qy    296 SSRVTTRHSFATLSFVAETFIFLYVGMD-----ALDIEK------------------W- 330
          || :   :  :|   : |   | | |:      || : |                    |
Db    345 SSLLIGLYLQST---IYEKSSFAYFGLAIKFAVALGLHKNSDDPSLTQNSKELRNRLLWS 401

Qy    331 -----KFVKNSQG---------LSVAVSSILVGLILVGRAAFVFPLSFLSNLAKKNSS-- 374
               :||  : |         :|:    ||  | : |  : |  :  : |||| :
Db    402 VFCIDRFVSMTTGRRPSIPLECISIPYPVILPDLEIPGSQSIVENMRAVINLAKLTNEIC 461

Qy    375 DKI------SFRQQI----------IIWWAGLMRGAV--SIALAYNKFTTSGHTS-LHEN 415
          | :      ||  |:          :| : |    |    |: :  | :: |   :::|
Db    462 DSLYWNPSPSFESQVNSVRRIYARLELWKSDLHSSVVFDESAVQHPLFRSNAHVQMIYDN 521

Qy    416 AIMITSTVTVV 426
          |||:|: | :|
Db    522 AIMLTTRVIMV 532


RESULT 9
US-08-928-692-13
; Sequence 13, Application US/08928692
; Patent No. 5958727
;  GENERAL INFORMATION:
;    APPLICANT:  Brody, Howard
;    APPLICANT:  Yaver, Deborah S.
;    APPLICANT:  Lamsa, Michael
;    APPLICANT:  Hansen, Kim
;    TITLE OF INVENTION:  Methods for Modifying the Production of
;    TITLE OF INVENTION:  a Polypeptide
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5958727o No. 5958727disk of No. 5958727th America, Inc.
;      STREET:  405 Lexington Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10174
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,692
;      FILING DATE:  12-SEPT-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4944.200-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  584 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 5958727e
US-08-928-692-13

  Query Match             3.6%;  Score 99;  DB 2;  Length 584;
  Best Local Similarity   18.7%;  Pred. No. 0.26;
  Matches   89;  Conservative   70;  Mismatches  172;  Indels  144;  Gaps   18;

Qy     73 SSHLLVFSEDLF-FIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIISFGA 131
          :| :: |:: |  ||  |  :   :||                :||     || ::|| |
Db     72 TSVMITFNQSLSPFIITLTFVASISGF--------------MFGYDTGYISSALISIGT 116

Qy    132 VKIFKHLDIDFLDFGD------YLAIGAIFAATDSVCTLQVLSQDETPLLYSLVFGEGVV 185
                ||   | :|:        ::||:  :  :     :  :    :  :|:|  | :
Db    117 -----DLDHKVLTYGEKEIVTAATSLGALITSIFAGTAADIFGRKRCLMGSNLMFVIGAI 171

Qy    186 NDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFG 245
                    :|    | :  :|  :|          |||| |    |  |:
Db    172 ----------LQVSAHTFWQMAVGRLIMG----------FGVGIGSL----IAPLFIS 205

Qy    246 RHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYT--WHNVTESSRVTTRH 303
                |:|  |:   |: ::  |:   | |  : ||  :::    |  :   | : |
Db    206 ------EIAPKMIRGRLT-VINSLWLTGGQLVAYGCGAGLNYVNNGWRILVGLSLIPTAV 258

Qy    304 SFATLSFVAETFIFLYVGMD---ALDIEK------------------------------ 329
           |  | |: :|  :  :  |   | :: |
Db    259 QFTCLCFLPDTPRYYVMKGDLARATEVLKRSYTDTSEEIIERKVEELVTLNQSIPGKNVP 318

Qy    330 ---WKFVKNSQGLSVAVSSILVGLILVGRAAFV-------FPLSFLSNLAKKNSS----- 374
             |  :|    :   : ::::|  |       |       |  :    :  ||||
Db    319 EKVWNTIKELHTVPSNLRALIIGCGLQAIQQFTGWNSLMYFSGTIFETVGFKNSSAVSII 378

Qy    375 ---------------DKISFRQQIIIWWAGLMRGAVSIALAYN----KF-------TTS 407
                          |||  |  ::|   |:    |  ::|::    ||        :|
Db    379 VSGTNFIFTLVAFFSIDKIGRRTILLIGLPGMTMALVVCSIAFHFLGIKFDGAVAVVVSS 438

Qy    408 GHTSLHENAIMITSTVTVVLFST-VVFGLMTKPLINLLLPPHKQMPSGHSSMTTS 461
          | :|       |   | :::|:     |: | |     | |      | |  | :
Db    439 GFSSWG-----IVIIVFIIVFAAFYALGIGTVPWQQSELFPQNVRGIGTSYATAT 488


RESULT 10
US-08-466-343D-2
; Sequence 2, Application US/08466343D
; Patent No. 6025154
;  GENERAL INFORMATION:
;    APPLICANT:  LI, Yi
;    TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN
```

```
;    TITLE OF INVENTION:  CHEMOKINE RECEPTOR HDGNR10 (AS AMENDED)
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;      STREET:  1100 NEW YORK AVE., NW, SUITE 600
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/466,343D
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEFFE, ERIC K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.1150000/EKS/KLM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  352 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-466-343D-2

  Query Match             3.5%;  Score 97.5;  DB 3;  Length 352;
  Best Local Similarity   21.8%;  Pred. No. 0.17;
  Matches   66;  Conservative   53;  Mismatches  127;  Indels   57;  Gaps   14;

Qy     173 PLLYSLVFGEGVV-NDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGI 231
           | ||||||  | | |    ::| |  :   ||               | | :|  |  : :
Db      34 PPLYSLVFIFGFVGNMLVILILINCQRLESMTDI----------YLLNLAISDLFFLLTV 83

Qy     232 GLLCAY------------IIKKLYF-GRHSTDREVALMMLMSYLSYIMAELFYLSGILT 277
                |             ::  ||| |  |    : |: :  ||: : | :| |
Db      84 PFWAHYAAAQWDFGNTMCQLLTGLYFIGFFSGIFFIILLTIDRYLAIVHA-VFALKARTV 142

Qy     278 VFFCGIVMSHYTWHNVTESSR---VTTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVK 334
           |  |:| |  ||     :|    : ||     | :   :  | |         :::| |
Db     143 TF--GVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTCSSH-FPY--------SQYQFWK 191

Qy     335 NSQGLSVAVSSILVGLI--LVGRAAFVFPLSFLSNLAKKNSSDKISFRQQII--IWWAGL 390
           | | | : :  ::: |:  ::  :  :  |    |  |:: : :: |  |:  ::||
Db     192 NFQTLKIVILGLVLPLLVMVICYSGILKTLLRCRNEKKRHRAVRLIFTIMIVYFLFWA-- 249

Qy     391 MRGAVSIALAYNKF------TTSGHTSLHENAIMITST--VTVVLFSTVVFGLMTKPLIN 442
              :| |  | |           ::  : |: :| |  :|    : :::  : :   |
Db     250 ---PYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQVTETLGMTHCCINPIIYAFVGEKFRN 306

Qy     443 LLL 445
           ||
Db     307 YLL 309


RESULT 11
US-09-087-232A-13
; Sequence 13, Application US/09087232A
; Patent No. 6153431
;  GENERAL INFORMATION:
;    APPLICANT:  Quillent et al.
;    TITLE OF INVENTION:  HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR
;    TITLE OF INVENTION:  VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION.
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Baker & Botts, L.L.P. attn. Lisa Kole
;      STREET:  30 Rockefeller Plaza
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10112
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/087,232A
;      FILING DATE:  28 MAY 1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/048,057
;      FILING DATE:  30 MAY 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KOLE, LISA B.
;      REGISTRATION NUMBER:  35,225
;      REFERENCE/DOCKET NUMBER:  AP 31115
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 408-2628
;      TELEFAX:  (212) 765-2519
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 352 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-087-232A-13

  Query Match             3.5%;  Score 97.5;  DB 4;  Length 352;
  Best Local Similarity   21.8%;  Pred. No. 0.17;
  Matches   66;  Conservative   53;  Mismatches  127;  Indels   57;  Gaps   14;

Qy     173 PLLYSLVFGEGVV-NDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGI 231
           | ||||||  | | |    ::| |  :   ||               | | :|  |  : :
Db      34 PPLYSLVFIFGFVGNMLVILILINCKRLKSMTDI----------YLLNLAISDLFFLLTV 83

Qy     232 GLLCAY------------IIKKLYF-GRHSTDREVALMMLMSYLSYIMAELFYLSGILT 277
                |             ::  ||| |  |    : |: :  ||: : | :| |
Db      84 PFWAHYAAAQWDFGNTMCQLLTGLYFIGFFSGIFFIILLTIDRYLAVVHA-VFALKARTV 142

Qy     278 VFFCGIVMSHYTWHNVTESSR---VTTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVK 334
           |  |:| |  ||     :|    : ||     | :   :  | |         :::| |
Db     143 TF--GVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTCSSH-FPY--------SQYQFWK 191

Qy     335 NSQGLSVAVSSILVGLI--LVGRAAFVFPLSFLSNLAKKNSSDKISFRQQII--IWWAGL 390
           | | | : :  ::: |:  ::  :  :  |    |  |:: : :: |  |:  ::||
Db     192 NFQTLKIVILGLVLPLLVMVICYSGILKTLLRCRNEKKRHRAVRLIFTIMIVYFLFWA-- 249

Qy     391 MRGAVSIALAYNKF------TTSGHTSLHENAIMITST--VTVVLFSTVVFGLMTKPLIN 442
              :| |  | |           ::  : |: :| |  :|    : :::  : :   |
Db     250 ---PYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQVTETLGMTHCCINPIIYAFVGEKFRN 306

Qy     443 LLL 445
           ||
Db     307 YLL 309


RESULT 12
US-08-861-105-14
; Sequence 14, Application US/08861105
; Patent No. 6258527
;  GENERAL INFORMATION:
;    APPLICANT:  LITTMAN, DAN R.
;    APPLICANT:  DENG, HONGKUI
;    APPLICANT:  ELLMEIER, WILFRIED
;    APPLICANT:  LANDAU, NATHANIEL R.
;    APPLICANT:  LIU, RONG
;    TITLE OF INVENTION:  G-COUPLED RECEPTORS ASSOCIATED WITH
;    TITLE OF INVENTION:  MACROPHAGE-TROPHIC HIV, AND DIAGNOSTIC AND THERAPEUTIC
```

```
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David A. Jackson, Esq.
;      STREET:  411 Hackensack Ave, Continental Plaza, 4th
;      STREET:  Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/861,105
;      FILING DATE:
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/666,020
;      FILING DATE:  19-JUN-1996
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/227,319
;      FILING DATE:  13-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  1049-1-004 N1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  352 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
US-08-861-105-14

  Query Match             3.5%;  Score 97.5;  DB 4;  Length 352;
  Best Local Similarity   21.8%;  Pred. No. 0.17;
  Matches   66;  Conservative   53;  Mismatches  127;  Indels   57;  Gaps   14;

Qy     173 PLLYSLVFGEGVV-NDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGI 231
           | ||||||  | | |    ::| |  :   ||              | | :|  |  : :
Db      34 PPLYSLVFIFGFVGNMLVILILINCKRLKSMTDI----------YLLNLAISDLFFLLTV 83

Qy     232 GLLCAY------------IIKKLYF-GRHSTDREVALMMLMSYLSYIMAELFYLSGILT 277
                |              ::  ||| |  |    : |: :  ||: : | :| |
Db      84 PFWAHYAAAQWDFGNTMCQLLTGLYFIGFFSGIFFIILLTIDRYLAVVHA-VFALKARTV 142

Qy     278 VFFCGIVMSHYTWHNVTESSR---VTTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVK 334
           |  |:| |  ||     :|    : ||     | :   :  | |          :::| |
Db     143 TF--GVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTCSSH-FPY--------SQYQFWK 191

Qy     335 NSQGLSVAVSSILVGLI--LVGRAAFVFPLSFLSNLAKKNSSDKISFRQQII--IWWAGL 390
           | | | : :  ::: |:  ::  :  :  |    |  |:: : :: |   |:  ::||
Db     192 NFQTLKIVILGLVLPLLVMVICYSGILKTLLRCRNEKKRHRAVRLIFTIMIVYFLFWA-- 249

Qy     391 MRGAVSIALAYNKF------TTSGHTSLHENAIMITST--VTVVLFSTVVFGLMTKPLIN 442
             :| |  | |          ::  : |: :| |  :|    : :::  : :   |
Db     250 ---PYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQVTETLGMTHCCINPIIYAFVGEKFRN 306

Qy     443 LLL 445
           ||
Db     307 YLL 309
```

```
RESULT 13
US-08-575-967A-2
; Sequence 2, Application US/08575967A
; Patent No. 6265184
;  GENERAL INFORMATION:
;    APPLICANT:  Gray et al.
;    TITLE OF INVENTION:  Chemokine Receptor Materials and Methods
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 S. Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/575,967A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6265184and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  32918
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-485-1900
;      TELEFAX:  206-485-1662
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  352 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  /= "88C amino acid sequence"
US-08-575-967A-2

  Query Match             3.5%;  Score 97.5;  DB 4;  Length 352;
  Best Local Similarity   21.8%;  Pred. No. 0.17;
  Matches   66;  Conservative   53;  Mismatches  127;  Indels   57;  Gaps   14;

Qy     173 PLLYSLVFGEGVV-NDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGI 231
           | ||||||  | | |    ::| |  :   ||              | | :|  |  : :
Db      34 PPLYSLVFIFGFVGNMLVILILINCKRLKSMTDI----------YLLNLAISDLFFLLTV 83

Qy     232 GLLCAY------------IIKKLYF-GRHSTDREVALMMLMSYLSYIMAELFYLSGILT 277
                |              ::  ||| |  |    : |: :  ||: : | :| |
Db      84 PFWAHYAAAQWDFGNTMCQLLTGLYFIGFFSGIFFIILLTIDRYLAVVHA-VFALKARTV 142

Qy     278 VFFCGIVMSHYTWHNVTESSR---VTTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVK 334
           |  |:| |  ||     :|    : ||     | :   :  | |          :::| |
Db     143 TF--GVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTCSSH-FPY--------SQYQFWK 191

Qy     335 NSQGLSVAVSSILVGLI--LVGRAAFVFPLSFLSNLAKKNSSDKISFRQQII--IWWAGL 390
           | | | : :  ::: |:  ::  :  :  |    |  |:: : :: |   |:  ::||
Db     192 NFQTLKIVILGLVLPLLVMVICYSGILKTLLRCRNEKKRHRAVRLIFTIMIVYFLFWA-- 249

Qy     391 MRGAVSIALAYNKF------TTSGHTSLHENAIMITST--VTVVLFSTVVFGLMTKPLIN 442
             :| |  | |          ::  : |: :| |  :|    : :::  : :   |
Db     250 ---PYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQVTETLGMTHCCINPIIYAFVGEKFRN 306

Qy     443 LLL 445
           ||
Db     307 YLL 309


RESULT 14
```

```
US-09-045-583-52
; Sequence 52, Application US/09045583
; Patent No. 6287805
;  GENERAL INFORMATION:
;    APPLICANT:  Graham, Gerard J. et al.
;    TITLE OF INVENTION:  No. 6287805el Molecules of the G Protein-Coupled
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/045,583
;      FILING DATE:  20-MAR-98
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E.
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  MNI-044
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)742-4214
;  INFORMATION FOR SEQ ID NO:  52:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  352 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  internal
US-09-045-583-52

  Query Match             3.5%;  Score 97.5;  DB 4;  Length 352;
  Best Local Similarity   21.8%;  Pred. No. 0.17;
  Matches   66;  Conservative   53;  Mismatches  127;  Indels   57;  Gaps   14;

Qy    173 PLLYSLVFGEGVV-NDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGI 231
          | ||||||  | | |    ::| |  :   ||              | | :|  |  : :
Db     34 PPLYSLVFIFGFVGNMLVILILINCKRLKSMTDI----------YLLNLAISDLFFLLTV 83

Qy    232 GLLCAY------------IIKKLYF-GRHSTDREVALMMLMSYLSYIMAELFYLSGILT 277
               |             ::  ||| |  |    : |: :  ||: : | :| |
Db     84 PFWAHYAAAQWDFGNTMCQLLTGLYFIGFFSGIFFIILLTIDRYLAIVHA-VFALKARTV 142

Qy    278 VFFCGIVMSHYTWHNVTESSR---VTTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVK 334
           |  |:| |  ||     :|    : ||     | :   :  | |          :::| |
Db    143 TF--GVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTCSSH-FPY--------SQYQFWK 191

Qy    335 NSQGLSVAVSSILVGLI--LVGRAAFVFPLSFLSNLAKKNSSDKISFRQQII--IWWAGL 390
          | | | : :  ::: |:  ::  :  :  |    |  |:: : :: |   |:  ::||
Db    192 NFQTLKIVILGLVLPLLVMVICYSGILKTLLRCRNEKKRHRAVRLIFTIMIVYFLFWA-- 249

Qy    391 MRGAVSIALAYNKF------TTSGHTSLHENAIMITST--VTVVLFSTVVFGLMTKPLIN 442
              :| |  | |           ::  : |: :| |  :|    : :::  : :   |
Db    250 ---PYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQVTETLGMTHCCINPIIYAFVGEKFRN 306

Qy    443 LLL 445
            ||
Db    307 YLL 309


RESULT 15
US-09-268-140-2
; Sequence 2, Application US/09268140
; Patent No. 6268176
; GENERAL INFORMATION:
;  APPLICANT:  Gemmill, Robert M.
;  APPLICANT:  Drabkin, Harry A.
;  TITLE OF INVENTION:  TRC8, A GENE RELATED TO THE HEDGEHOG RECEPTOR, PATCHED
;  FILE REFERENCE:  93445-00004
;  CURRENT APPLICATION NUMBER: US/09/268,140
;  CURRENT FILING DATE:  2000-03-12
;  PRIOR APPLICATION NUMBER:  US 60/077,723
;  PRIOR FILING DATE:  1998-03-12
;  NUMBER OF SEQ ID NOS:  46
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 664
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-268-140-2

  Query Match             3.5%;  Score 96;  DB 4;  Length 664;
  Best Local Similarity   17.4%;  Pred. No. 0.63;
  Matches   83;  Conservative   87;  Mismatches  182;  Indels  126;  Gaps   19;

Qy      1 MAFGLS--SLLQNSDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESI---TAL 55
            : ||:   :|||      | :: :: ::| |       |:| : |    : |::   ::|
Db    135 LTFGIGYVTLLQ----IHSIYSQLIILDLLVP------VIGLITELPLHIRETLLFTSSL 184

Qy     56 IIGLCTGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNA-GFQVKKKQFF--VNFMTI 112
          |: | |  |: :        | |     ::|||   ::   | |:  :  :  : | |
Db    185 ILTLNTVFVLAV-------KLKWFYYSTRYVYLLVRHMYRIYGLQLLMEDTWKRIRFPDI 237

Qy    113 MLFGAIGTLISCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDET 172
          :   :  : : : :   :::    |  |: : |:         | :| | :   | |
Db    238 LRVFWLTRVTAQATVLMYILRMANETDSFFISWDDF---------WDLICNLIISGCDST 288

Qy    173 PLLYSLVFGEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGI- 231
              | |   |  : :  |   | :|  :: :    | |:   |:  |:   |  |:
Db    289 ----LTVLGMSAVISSVAHYLGLGILAFIGSTEEDDRRLGFVAPVLFFILA---LQTGLS 341

Qy    232 GL-----------LCAYIIKKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVF 279
            ||           :|  :   |:|    ||  | : :  |::|        |     :|
Db    342 GLRPEERLIRLSRNMCLLLTAVLHFIHGMTD-PVLMSLSASHVSSFRRHFPVLFVSACLF 400

Qy    280 FCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETFIFLYVGMDAL------------- 325
          :::|:  ||:   :: :     ||  :|  |  : : | :
Db    401 ILPVLLSYVLWHHYALNTWL-----FAVTAFCVELCLKVIVSLTVYTLFMIDGYYNVLWE 455

Qy    326 DIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLS--------------FLSNLAK 370
           :: : :   | |   ::   : |::: |  |:                    |   |
Db    456 KLDDYVYYVRSTG---SIIEFIFGVVMFGNGAYTMMFESGSKIRAFMMCLHAYFNIYLQA 512

Qy    371 KNSSDKISFRQQIIIWWAGLMRGAVSI--------------------ALAYNKFTTS 407
          ||              |   |   ::                     |: |::|||
Db    513 KNG-------------WKTFMNRRTAVKKINSLPEIKGSRLQEINDVCAICYHEFTTS 557


Search completed: October 18, 2002, 12:24:38
Job time : 19 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: October 18, 2002, 12:24:11 ; Search time 31 Seconds
(without alignments)
3024.621 Million cell updates/sec

Title: US-09-830-123-2
Perfect score: 2768
Sequence: 1 MAFGLSSLLQNSDLFTSDHA.....FGGRGFVPFVAGSPVEQSPR 542

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2768	100.0	542	10	Q9FEB3	Q9feb3 pharbitis n
2	2726	98.5	542	10	Q94IE0	Q94ie0 ipomoea tri
3	2127.5	76.9	555	10	Q9FZN0	Q9fzn0 atriplex gm
4	2124.5	76.8	553	10	Q94LX4	Q94lx4 nierembergi
5	2103.5	76.0	540	10	Q94LX5	Q94lx5 petunia hyb
6	2083.5	75.3	542	10	Q9ARH6	Q9arh6 citrus para
7	2072	74.9	556	10	Q94K25	Q94k25 suaeda mari
8	2055.5	74.3	535	10	Q9SXJ8	Q9sxj8 oryza sativ
9	2054.5	74.2	546	10	Q9CAW6	Q9caw6 arabidopsis
10	2041	73.7	538	10	Q9ZPK3	Q9zpk3 arabidopsis
11	1953.5	70.6	555	10	Q94LX3	Q94lx3 torenia hyb
12	1925.5	69.6	552	10	Q9SQU0	Q9squ0 arabidopsis
13	1865.5	67.4	546	10	Q94BM4	Q94bm4 triticum ae
14	1819.5	65.7	534	10	Q93YH2	Q93yh2 lycopersico
15	1732.5	62.6	457	10	O04655	O04655 arabidopsis
16	1600.5	57.8	529	10	Q9FJ63	Q9fj63 arabidopsis
17	1107.5	40.0	295	10	Q94BM3	Q94bm3 triticum ae
18	620	22.4	561	5	Q9XZH4	Q9xzh4 drosophila
19	620	22.4	649	5	Q9VPJ1	Q9vpj1 drosophila
20	591.5	21.4	703	5	Q9XW14	Q9xw14 caenorhabdi
21	585.5	21.2	725	4	Q96T83	Q96t83 homo sapien
22	576.5	20.8	687	5	Q9U624	Q9u624 drosophila
23	575.5	20.8	727	5	Q9VM99	Q9vm99 drosophila
24	563.5	20.4	616	4	O75827	O75827 homo sapien
25	556	20.1	569	3	O13726	O13726 schizosacch
26	552.5	20.0	629	5	Q20944	Q20944 caenorhabdi
27	526	19.0	153	10	Q9LKH6	Q9lkh6 mesembryant
28	523	18.9	140	10	Q9LKH5	Q9lkh5 mesembryant
29	512.5	18.5	531	10	Q93YH1	Q93yh1 lycopersico
30	483	17.4	518	5	Q9VIF9	Q9vif9 drosophila
31	474.5	17.1	1203	5	Q9NGZ4	Q9ngz4 drosophila
32	471	17.0	1179	5	Q9NCQ0	Q9ncq0 aedes aegyp
33	469.5	17.0	560	3	Q9HEX3	Q9hex3 pneumocysti
34	451	16.3	813	13	Q9W6Q1	Q9w6q1 amphiuma tr
35	448	16.2	437	4	Q9Y507	Q9y507 homo sapien
36	445.5	16.1	698	13	Q9W724	Q9w724 cyprinus ca
37	442	16.0	673	5	Q23706	Q23706 carcinus ma
38	442	16.0	779	13	Q9W714	Q9w714 platichthys
39	441.5	16.0	781	13	P70009	P70009 xenopus lae
40	425.5	15.4	555	4	Q96EM2	Q96em2 homo sapien
41	425.5	15.4	634	5	Q9TZH0	Q9tzh0 caenorhabdi
42	421	15.2	106	10	Q9ATZ9	Q9atz9 zea mays (m
43	419	15.1	478	5	O01995	O01995 caenorhabdi
44	410	14.8	411	4	Q9Y2E8	Q9y2e8 homo sapien
45	409.5	14.8	627	5	Q09432	Q09432 caenorhabdi

ALIGNMENTS

RESULT 1
Q9FEB3
ID Q9FEB3 PRELIMINARY; PRT; 542 AA.
AC Q9FEB3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NA+/H+ EXCHANGER (NA+/H+ EXCHANGER PROTEIN).
GN PURPLE OR INNHX1.
OS Pharbitis nil (Violet) (Japanese morning glory).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Convolvulaceae; Ipomoea.
OX NCBI_TaxID=35883;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KKZSK-2; TISSUE=LEAF, AND FULLY COLORED FLOWER BUDS;
RX MEDLINE=20487008; PubMed=11034195;
RA Fukada-Tanaka S., Inagaki Y., Yamaguchi T., Saito N., Iida S.;
RT "Colour-enhancing protein in blue petals.";
RL Nature 407:581-581(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PR-R; TISSUE=LEAF;
RX MEDLINE=21276661; PubMed=11382810;
RA Yamaguchi T., Fukada-Tanaka S., Inagaki Y., Saito N.,
RA Yonekura-Sakakibara K., Tanaka Y., Kusumi T., Iida S.;
RT "Genes Encoding the Vacuolar Na+/H+ Exchanger and Flower Coloration.";
RL Plant Cell Physiol. 42:451-461(2001).
DR EMBL; AB033990; BAB16381.1; -.
DR EMBL; AB033989; BAB16380.1; -.
DR EMBL; AB055062; BAB60899.1; -.
DR InterPro; IPR001179; FKBP_PPIase.
DR InterPro; IPR000676; NaH_Exchngr.
DR Pfam; PF00999; Na_H_Exchanger; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
DR PROSITE; PS00453; FKBP_PPIASE_1; UNKNOWN_1.
SQ SEQUENCE 542 AA; 59973 MW; 4B47FDE04401A191 CRC64;

```
  Query Match           100.0%;  Score 2768;  DB 10;  Length 542;
  Best Local Similarity  100.0%;  Pred. No. 7.1e-204;
  Matches  542;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MAFGLSSLLQNSDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MAFGLSSLLQNSDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLC 60

Qy      61 TGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGT 120

Qy     121 LISCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVF 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LISCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVF 180

Qy     181 GEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIK 240

Qy     241 KLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 KLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVT 300

Qy     301 TRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVF 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVF 360

Qy     361 PLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMIT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 PLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMIT 420

Qy     421 STVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDS 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 STVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDS 480

Qy     481 ESDMITGPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVEQS 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ESDMITGPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVEQS 540

Qy     541 PR 542
           ||
Db     541 PR 542


RESULT 2
Q94IE0
ID   Q94IE0      PRELIMINARY;      PRT;   542 AA.
AC   Q94IE0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NA+/H+ EXCHANGER.
GN   ITNHX1.
OS   Ipomoea tricolor (Morning glory).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Convolvulaceae; Ipomoea.
OX   NCBI_TaxID=89664;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21276661; PubMed=11382810;
RA   Yamaguchi T., Fukada-Tanaka S., Inagaki Y., Saito N.,
RA   Yonekura-Sakakibara K., Tanaka Y., Kusumi T., Iida S.;
RT   "Genes Encoding the Vacuolar Na+/H+ Exchanger and Flower Coloration.";
RL   Plant Cell Physiol. 42:451-461(2001).
DR   EMBL; AB054979; BAB60901.1; -.
SQ   SEQUENCE   542 AA;  60024 MW;  0C8381CFB78B239C CRC64;

  Query Match           98.5%;  Score 2726;  DB 10;  Length 542;
  Best Local Similarity  98.7%;  Pred. No. 1.2e-200;
  Matches  535;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 MAFGLSSLLQNSDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLC 60
           ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MAFGLSSLLQNSELFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLC 60

Qy      61 TGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGT 120
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TGVVILLLSRGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGT 120

Qy     121 LISCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVF 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LISCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVF 180

Qy     181 GEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIK 240

Qy     241 KLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 KLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVT 300

Qy     301 TRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVF 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVF 360

Qy     361 PLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMIT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 PLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMIT 420

Qy     421 STVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDS 480
           ||||||||||||||||||||||||||||||: |||||||||||||||| ||||||| ||
Db     421 STVTVVLFSTVVFGLMTKPLINLLLPPHKQIASGHSSMTTSEPSSPKHFAVPLLDNQHDS 480

Qy     481 ESDMITGPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVEQS 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     481 ESDMITGPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPAEQS 540

Qy     541 PR 542
           ||
Db     541 PR 542


RESULT 3
Q9FZN0
ID   Q9FZN0      PRELIMINARY;      PRT;   555 AA.
AC   Q9FZN0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NA/H ANTIPORTER NHX1.
GN   AGNHX1.
OS   Atriplex gmelini.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllidae; Caryophyllales; Chenopodiaceae; Atriplex.
OX   NCBI_TaxID=118074;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21330247; PubMed=11437248;
RA   Hamada A., Shono M., Xia T., Ohta M., Hayashi Y., Tanaka A.,
RA   Hayakawa T.;
RT   "Isolation and characterization of a Na+/H+ antiporter gene from the
RT   halophyte Atriplex gmelini.";
RL   Plant Mol. Biol. 46:35-42(2001).
DR   EMBL; AB038492; BAB11940.1; -.
DR   InterPro; IPR001179; FKBP_PPIase.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
DR   PROSITE; PS00453; FKBP_PPIASE_1; UNKNOWN_1.
SQ   SEQUENCE   555 AA;  61504 MW;  ACD5ED45FF3D398A CRC64;

  Query Match           76.9%;  Score 2127.5;  DB 10;  Length 555;
  Best Local Similarity  76.6%;  Pred. No. 7.8e-155;
```

```
  Matches  418;  Conservative   50;  Mismatches   65;  Indels   13;  Gaps    3;

Qy       5 LSSLLQN--SDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLCTG 62
           |||||      | ||||||||||||||||| |||:|||||||||:|||||||:||| ||
Db       5 LSSLLSGKMDALTTSDHASVVSMNLFVALLCGCIVIGHLLEENRWMNESITALLIGLATG 64

Qy      63 VVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLI 122
           |||||:|||||||||||||||||||||||||||||||||||| ||:||:||||:|||:
Db      65 VVILLISGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFRNFITIVLFGAVGTLV 124

Qy     123 SCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVFGE 182
           | :||| ||: ||| ||| |:  |||||||||||||||||||:||||||||||||||
Db     125 SFTIISLGALSIFKKLDIGTLELADYLAIGAIFAATDSVCTLQVLNQDETPLLYSLVFGE 184

Qy     183 GVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKL 242
           |||||||||||||||||:|  | :| | |:||||||::|| ||   ||| |||||||
Db     185 GVVNDATSVVLFNAIQSFDLTRIDHRIALQFMGNFLYLFIASTILGAFTGLLSAYIIKKL 244

Qy     243 YFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTR 302
           ||||||||||||||||:||||::|||||||||||||||||||||||||||||||||||:
Db     245 YFGRHSTDREVALMMLMAYLSYMLAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTK 304

Qy     303 HSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPL 362
           |:||||||||| |:|||||||||||||:|| :| |:||||||||::|||:|||||||||
Db     305 HAFATLSFVAEVFLFLYVGMDALDIEKWRFVSDSPGISVAVSSILLGLVMVGRAAFVFPL 364

Qy     363 SFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITST 422
           |:| | |||: |:|::| |||:|||||||||||:|||||:|| |||| |  ||||||||
Db     365 SWLMNFAKKSQSEKVTFNQQIVIWWAGLMRGAVSMALAYNQFTRSGHTQLRGNAIMITST 424

Qy     423 VTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSES 482
           ::||||||:||||:|||||  |||  |       |  | |:  ||| :::|||:   | |
Db     425 ISVVLFSTMVFGLLTKPLIMFLLPQPKHFT---SCSTVSDVGSPKSYSLPLLEGNQDYEV 481

Qy     483 DMITGPE--------VARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAG 534
           |:  |          : ||::|||||  |||||| ||||||||||||||||||||| |
Db     482 DVGNGNHEDTTEPRTIVRPSSLRMLLNAPTHTVHHYWRKFDDSFMRPVFGGRGFVPFVPG 541

Qy     535 SPVEQS 540
           || |||
Db     542 SPTEQS 547


RESULT 4
Q94LX4
ID   Q94LX4      PRELIMINARY;      PRT;   553 AA.
AC   Q94LX4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NA H-ANTIPORTOR.
GN   NHPURPLE.
OS   Nierembergia caerulea.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Nierembergia.
OX   NCBI_TaxID=144865;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Iida S., Kusumi T., Yonekura-Sakakibara K., Tanaka Y.;
RT   "Plant Na-H antiportor.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB051818; BAB56106.1; -.
SQ   SEQUENCE   553 AA;  61119 MW;  9F2B47A49D7AB2CC CRC64;

  Query Match           76.8%;  Score 2124.5;  DB 10;  Length 553;
  Best Local Similarity   77.0%;  Pred. No. 1.3e-154;
  Matches  415;  Conservative   51;  Mismatches   66;  Indels    7;  Gaps    3;

Qy       1 MAFGLSSLL-QNSDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGL 59
           |||   :|| : ::| |||| ||||:||||||:|||||:|||||||||:|||||||:||
Db       1 MAFDFGTLLGKMNNLTTSDHQSVVSVNLFVALICACIVIGHLLEENRWMNESITALVIGS 60

Qy      60 CTGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIG 119
           ||||:|||:||||:||:|||||||||||||||||||||||||| || || |||||||:|
Db      61 CTGVIILLISGGKNSHILVFSEDLFFIYLLPPIIFNAGFQVKKKSFFRNFSTIMLFGAVG 120

Qy     120 TLISCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLV 179
           ||||  ||| ||: ||| :||  |: |||||||||||||||||||:|:|||||||||
Db     121 TLISFIIISAGAIGIFKKMDIGHLEIGDYLAIGAIFAATDSVCTLQVLNQEETPLLYSLV 180

Qy     180 FGEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYII 239
           |||||||||||||||||:|:||::       |  |||||||| |||||| :||| |:||
Db     181 FGEGVVNDATSVVLFNAVQNFDLSHISTGKALQLIGNFLYLFASSTFLGVAVGLLSAFII 240

Qy     240 KKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRV 299
           |||||||||||||||:|:||:||||::|||||||||||||||||||||||||||||||||
Db     241 KKLYFGRHSTDREVAIMILMAYLSYMLAELFYLSGILTVFFCGIVMSHYTWHNVTESSRV 300

Qy     300 TTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFV 359
           ||:|:|||||||:|| |||||||||||||||||| :| | |: |||||:||:|||| |||
Db     301 TTKHTFATLSFIAEIFIFLYVGMDALDIEKWKFVSDSPGTSIKVSSILLGLVLVGRGAFV 360

Qy     360 FPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMI 419
           ||||||||| |||  ||||| ||: ||||||||||||:|||||:||  ||| |  |||||
Db     361 FPLSFLSNLTKKNPEDKISFNQQVTIWWAGLMRGAVSMALAYNQFTRGGHTQLRANAIMI 420

Qy     420 TSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPD 479
           |||:|||||||||||||||| |||  | :       | :||| :|| | |||||:  |
Db     421 TSTITVVLFSTVVFGLMTKPLILLLLLPSQKHL----IRMISSEPMTPKSFIVPLLDSTQD 476

Qy     480 SESDMITGPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVE 538
           ||:|:  |  | || :||||| ||:|||| ||||||::|||||||||||||| ||| |
Db     477 SEADL--GRHVPRPHSLRMLLSTPSHTVHYYWRKFDNAFMRPVFGGRGFVPFVPGSPTE 533


RESULT 5
Q94LX5
ID   Q94LX5      PRELIMINARY;      PRT;   540 AA.
AC   Q94LX5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NA H-ANTIPORTOR.
GN   PHPURPLE.
OS   Petunia hybrida (Petunia).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Petunia.
OX   NCBI_TaxID=4102;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Iida S., Kusumi T., Yonekura-Sakakibara K., Tanaka Y.;
RT   "Plant Na-H antiportor.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB051817; BAB56105.1; -.
SQ   SEQUENCE   540 AA;  59510 MW;  BCE2740F275E896A CRC64;

  Query Match           76.0%;  Score 2103.5;  DB 10;  Length 540;
  Best Local Similarity   76.4%;  Pred. No. 5.2e-153;
  Matches  412;  Conservative   51;  Mismatches   69;  Indels    7;  Gaps    3;

Qy       1 MAFGLSSLLQNSD-LFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGL 59
           |||   :|| | | | |||| ||||:||||||:|||||:|||||||||:|||||||:||
Db       1 MAFDFGTLLGNVDRLSTSDHQSVVSINLFVALICACIVIGHLLEENRWMNESITALVIGS 60

Qy      60 CTGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIG 119
           |||:||||:||||:||:|||||||||||||||||||||||||| || || |||||||:|
Db      61 CTGIVILLISGGKNSHILVFSEDLFFIYLLPPIIFNAGFQVKKKSFFRNFSTIMLFGALG 120

Qy     120 TLISCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLV 179
           ||||  ||| ||: ||| ::|  |: |||||||||:|||||||||:|||||||||
Db     121 TLISFIIISLGAIGIFKKMNIGSLEIGDYLAIGAIFSATDSVCTLQVLNQDETPLLYSLV 180

Qy     180 FGEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYII 239
           |||||||||||||||||:||::  |    :  :|||||| ||| |||  ||| ||||
```

```
Db    181 FGEGVVNDATSVVLFNAIQNFDLSHIDTGKAMELVGNFLYLFASSTALGVAAGLLSAYII 240

Qy    240 KKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRV 299

Db    241 KKLYFGRHSTDREVAIMILMAYLSYMLAELFYLSAILTVFFSGIVMSHYTWHNVTESSRV 300

Qy    300 TTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFV 359

Db    301 TTKHTFATLSFIAEIFIFLYVGMDALDIEKWKFVSDSPGISVQVSSILLGLVLVGRAAFV 360

Qy    360 FPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMI 419

Db    361 FPLSFLSNLTKKTPEAKISFNQQVTIWWAGLMRGAVSMALAYNQFTRGGHTQLRANAIMI 420

Qy    420 TSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPD 479

Db    421 TSTITVVLFSTVVFGLMTKPLIRILLPSHKHL----SRMISSEPTTPKSFIVPLLDSTQD 476

Qy    480 SESDMITGPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVE 538

Db    477 SEADL--ERHVPRPHSLRMLLSTPSHTVHYYWRKFDNAFMRPVFGGRGFVPFAPGSPTD 533


RESULT 6
Q9ARH6
ID   Q9ARH6      PRELIMINARY;      PRT;   542 AA.
AC   Q9ARH6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SODIUM/PROTON EXCHANGER.
OS   Citrus paradisi (Grapefruit).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Sapindales; Rutaceae; Citrus.
OX   NCBI_TaxID=37656;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PEEL;
RA   Porat R., Lurie S., Povancello D.;
RT   "A heat treatment induced the transcription of a sodium proton
RT   exchanger gene in grapefruit peel tissue.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY028416; AAK27314.2; -.
SQ   SEQUENCE   542 AA;  59836 MW;  1A8525F2C5FAEE2C CRC64;

  Query Match             75.3%;  Score 2083.5;  DB 10;  Length 542;
  Best Local Similarity   75.4%;  Pred. No. 1.8e-151;
  Matches  407;  Conservative   61;  Mismatches   65;  Indels    7;  Gaps    5;

Qy      5 LSSLLQNSDLF-TSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLCTGV 63

Db      5 ISSVVRKLQMVNTSDHNSVVSINIFVALPCASIVIGHLLEESRWMNESITALLIGVCAGV 64

Qy     64 VILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLIS 123

Db     65 IILLTTGGKSSHLFVFSEDLFFIYVLPPIIFNAGFQVKKKQFFRNFITIMLFGAIGTLVS 124

Qy    124 CSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVFGEG 183

Db    125 CTIISLGVIQFFKKLDIGTLDIGDYLAIGAIFAATDSVCTLQVLNQDDTPLLYSLVFGEG 184

Qy    184 VVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKLY 243

Db    185 VVNDATSVVLFNAIQSFDLTHINTRSAFQFIGNFLYLFFTSTLLGVIGGLLSAYVIKKLY 244

Qy    244 FGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRH 303

Db    245 FGRHSTDREVAIMVLMAYLSYMLAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTKH 304

Qy    304 SFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLS 363

Db    305 TFATLSFVAEIFTFLYVGMDALDIEKWRFVKGSPGTSVAASAMLMGLIMAGRAAFVFPLS 364

Qy    364 FLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTV 423

Db    365 FLTNLAKKSPTEKISIKQQVIIWWAGLMRGAVSMALAYNQFTRSGHTQLRENAIMITSTI 424

Qy    424 TVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLL-DNQPDSES 482

Db    425 TVVLFSTVVFGLMTEPLIRLLL-PHPKHTTNH---ILSDPSTPKSLSQPLLEEGQQDSYA 480

Qy    483 DMITGPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVEQSPR 542

Db    481 DLV-GPTVPRPGSLRALLTTPTHTVHYYWRKFDDAFMRPVFGGRGFAPFVPGSPTERSVR 539


RESULT 7
Q94K25
ID   Q94K25      PRELIMINARY;      PRT;   556 AA.
AC   Q94K25;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NA+/H+ ANTIPORTER.
OS   Suaeda maritima subsp. salsa.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllidae; Caryophyllales; Chenopodiaceae; Suaeda.
OX   NCBI_TaxID=126914;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhang H., Zhang Q., Ma X.;
RT   "Na+/H+ antiporter in Suaeda salsa.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF370358; AAK53432.1; -.
SQ   SEQUENCE   556 AA;  61672 MW;  DDF6AB696647D48E CRC64;

  Query Match             74.9%;  Score 2072;  DB 10;  Length 556;
  Best Local Similarity   75.6%;  Pred. No. 1.4e-150;
  Matches  402;  Conservative   58;  Mismatches   62;  Indels   10;  Gaps    2;

Qy     16 TSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLCTGVVILLLSGGKSSH 75

Db     18 TSDHASVVSMNLFVALLRGCIVIGHLLEENRWMNESITALLIGLSTGIIILLISGGKSSH 77

Qy     76 LLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIISFGAVKIF 135

Db     78 LLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFRNFITIILFGAVGTLVSFIIISLGSIAIF 137

Qy    136 KHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVFGEGVVNDATSVVLFN 195

Db    138 QKMDIGSLELGDLLAIGAIFAATDSVCTLQVLNQDETPLLYSLVFGEGVVNDATSVVLFN 197

Qy    196 AIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDREVAL 255

Db    198 AIQNFDLTHIDHRIAFQFGGNFLYLFFASTLLGAVTGLLSAYVIKKLYFGRHSTDREVAL 257

Qy    256 MMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETF 315

Db    258 MMLMAYLSYMLAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTKHAFATLSFVAEIF 317

Qy    316 IFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKKNSSD 375

Db    318 IFLYVGMDALDIEKWRFVSDSPGTSVAVSSILLGLHMVGRAAFVFPFAFLMNLSKKSNSE 377

Qy    376 KISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTVVFGL 435

Db    378 KVTFNQQIVIWWAGLMKSAVSVALAYNQFSRSGHTQLRGNAIMITSTITVVLFSTMVFGL 437

Qy    436 MTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMITGPE------ 489

Db    438 LTKPLILFMLPQPKHFT---SASTVSDLGSPKSFSLPLLEDRQDSEADLGNDDEEAYPRG 494

Qy    490 -VARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVEQS 540
```

```
Db    495 TIARPTSLRMLLNAPTHTVHHYWRRFDDYFMRPVFGGRGFVPFVPGSPTEQS 546


RESULT 8
Q9SXJ8
ID   Q9SXJ8      PRELIMINARY;      PRT;   535 AA.
AC   Q9SXJ8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OSNHX1 PROTEIN.
GN   OSNHX1.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. NIPPONBARE;
RX   MEDLINE=99326147; PubMed=10395929;
RA   Fukuda A., Nakamura A., Tanaka Y.;
RT   "Molecular cloning and expression of the Na+/H+ exchanger gene in
RT   Oryza sativa.";
RL   Biochim. Biophys. Acta 1446:149-155(1999).
DR   EMBL; AB021878; BAA83337.1; -.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
SQ   SEQUENCE   535 AA;  59070 MW;  6BCADDB29B131976 CRC64;

  Query Match             74.3%;  Score 2055.5;  DB 10;  Length 535;
  Best Local Similarity   76.4%;  Pred. No. 2.4e-149;
  Matches  401;  Conservative   49;  Mismatches   66;  Indels    9;  Gaps    3;

Qy     16 TSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLCTGVVILLLSGGKSSH 75
          |||:||||||:|||||||||||||||||||||||||||||||||||||||||||:: |||||
Db     16 TSDYASVVSINLFVALLCACIVLGHLLEENRWVNESITALIIGLCTGVVILLMTKGKSSH 75

Qy     76 LLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIISFGAVKIF 135
          | |||||||||||||||||||||||||||||| ||||| ||||:||:||   ||  |: ||
Db     76 LFVFSEDLFFIYLLPPIIFNAGFQVKKKQFFRNFMTITLFGAVGTMISFFTISIAAIAIF 135

Qy    136 KHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVFGEGVVNDATSVVLFN 195
            ::|  || ||:|||||||:|||||||||||:||||| |||||||||||||||||:||||
Db    136 SRMNIGTLDVGDFLAIGAIFSATDSVCTLQVLNQDETPFLYSLVFGEGVVNDATSIVLFN 195

Qy    196 AIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDREVAL 255
          |:|:||:   |  : | |:||| |||||||||||  ||| |||||||| ||||||||||
Db    196 ALQNFDLVHIDAAVVLKFLGNFFYLFLSSTFLGVFAGLLSAYIIKKLYIGRHSTDREVAL 255

Qy    256 MMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETF 315
          ||||:||||::|||  |||||||||||||||||||||||||||:|:||||||:||||
Db    256 MMLMAYLSYMLAELLDLSGILTVFFCGIVMSHYTWHNVTESSRVTTKHAFATLSFIAETF 315

Qy    316 IFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKKNSSD 375
          :|||||||||||||:|  :  | |: :||||:||:|:|||||||||||||||| ||  ::
Db    316 LFLYVGMDALDIEKWEFASDRPGKSIGISSILLGLVLIGRAAFVFPLSFLSNLTKKAPNE 375

Qy    376 KISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTVVFGL 435
          ||::|||::||||||||||||||||||| |||| || ||||||||:|||||||:|||:
Db    376 KITWRQQVVIWWAGLMRGAVSIALAYNKFTRSGHTQLHGNAIMITSTITVVLFSTMVFGM 435

Qy    436 MTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMITGPEVARPTA 495
          |||||| ||||      |||    ||||||||    ||| :   |: :  |   : ||::
Db    436 MTKPLIRLLLP-----ASGHP--VTSEPSSPKSLHSPLLTSMQGSDLESTT--NIVRPSS 486

Qy    496 LRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVEQS 540
          |||||  |||||| |||||||: |||:|||||||||  ||| |||     .
Db    487 LRMLLTKPTHTVHYYWRKFDDALMRPMFGGRGFVPFSPGSPTEQS 531


RESULT 9
```

```
Q9CAW6
ID   Q9CAW6      PRELIMINARY;      PRT;   546 AA.
AC   Q9CAW6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PUTATIVE SODIUM PROTON EXCHANGER.
GN   T9J14.2.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=21016720; PubMed=11130713;
RA   Salanoubat M., Lemcke K., Rieger M., Ansorge W., Unseld M.,
RA   Fartmann B., Valle G., Bloecker H., Perez-Alonso M., Obermaier B.,
RA   Delseny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA   De Simone V., Choisne N., Artiguenave F., Robert C., Brottier P.,
RA   Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,
RA   Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA   Wurmbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA   Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA   Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,
RA   Conrad A., Hornischer K., Kauer G., Loehnert T.-H., Nordsiek G.,
RA   Reichelt J., Scharfe M., Schoen O., Bargues M., Terol J., Climent J.,
RA   Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA   Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,
RA   de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA   Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,
RA   Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA   Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA   Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
RA   Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA   Pai G., Militscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA   Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA   Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA   Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA   Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA   Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT   "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 408:820-822(2000).
DR   EMBL; AC009465; AAG51408.1; -.
DR   InterPro; IPR001179; FKBP_PPIase.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
DR   PROSITE; PS00453; FKBP_PPIASE_1; UNKNOWN_1.
SQ   SEQUENCE   546 AA;  60522 MW;  BEA270D40446360B CRC64;

  Query Match             74.2%;  Score 2054.5;  DB 10;  Length 546;
  Best Local Similarity   75.4%;  Pred. No. 3e-149;
  Matches  401;  Conservative   56;  Mismatches   60;  Indels   15;  Gaps    4;

Qy     16 TSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLCTGVVILLLSGGKSSH 75
          |||||||||:|||||||||||||:|||||||||:|||||:||| |||||||:| ||:||
Db     16 TSDHASVVSLNLFVALLCACIVIGHLLEENRWMNESITALLIGLGTGVVILLISRGKNSH 75

Qy     76 LLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIISFGAVKIF 135
          |||||||||||||||||||||||||||||| ||:||| ||||||::||:||| ||:: |
Db     76 LLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFRNFVTIMAFGAIGTVVSCTIISLGAIQFF 135

Qy    136 KHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVFGEGVVNDATSVVLFN 195
          | |||   | ||:||||||||||||||||:|||||||||||||||||||||||||||||
Db    136 KKLDIGTFDLGDFLAIGAIFAATDSVCTLQVLNQDETPLLYSLVFGEGVVNDATSVVLFN 195

Qy    196 AIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDREVAL 255
          ||||||:|  : :    |:||| |||| || |||  ||: ||:||||||||||||||||||
Db    196 AIQSFDLTHLNHEAAFQFLGNFFYLFLLSTGLGVATGLISAYVIKKLYFGRHSTDREVAL 255
```

```
Qy    256 MMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFVAETF 315
          ||||:||||::|||| |||||||||||||||||||||||||||:||:|:||||||:||||
Db    256 MMLMAYLSYMLAELFALSGILTVFFCGIVMSHYTWHNVTESSRITTKHAFATLSFLAETF 315

Qy    316 IFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKKNSSD 375
          |||||||||||||||:|| :| | ||||||||:||:::|||||||||||||||||||: |:
Db    316 IFLYVGMDALDIEKWRFVSDSPGTSVAVSSILMGLVMLGRAAFVFPLSFLSNLAKKHQSE 375

Qy    376 KISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTVVFGL 435
          ||| :||::|||||||||||:|||||||| |||| |  ||||||||:|| ||||:|||:
Db    376 KISIKQQVVIWWAGLMRGAVSMALAYNKFTRSGHTELRGNAIMITSTITVCLFSTMVFGM 435

Qy    436 MTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQ-------PDSESDMITGP 488
          :|||||  |: ||::  :  :||  |: |:||   :|||| :       | |  |
Db    436 LTKPLIRYLM-PHQKATTSTTSM-LSDDSTPKSIHIPLLDGEQLDSFELPGSHQD----- 488

Qy    489 EVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVEQS 540
           | || :||  |  || ||| |||:|||:||||||||||||||| ||| |:|
Db    489 -VPRPNSLRGFLMRPTRTVHYYWRQFDDAFMRPVFGGRGFVPFVPGSPTERS 539


RESULT 10
Q9ZPK3
ID   Q9ZPK3         PRELIMINARY;   PRT;   538 AA.
AC   Q9ZPK3;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   SODIUM PROTON EXCHANGER NHX1 (FRAGMENT).
GN   NHE1.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=99145575; PubMed=9990049;
RA   Gaxiola R.A., Rao R., Sherman A., Grisafi P., Alper S.L., Fink G.R.;
RT   "The Arabidopsis thaliana proton transporters, AtNhx1 and Avp1, can
RT   function in cation detoxification in yeast.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:1480-1485(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Quintero F.J., Blatt M.R., Pardo J.M.;
RT   "The AtNHE1 gene encodes a putative Na+/H+ antiporter.";
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF106324; AAD16946.1; -.
DR   EMBL; AF056190; AAF21755.1; -.
DR   InterPro; IPR001179; FKBP_PPIase.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
DR   PROSITE; PS00453; FKBP_PPIASE_1; UNKNOWN_1.
FT   NON_TER     538    538
SQ   SEQUENCE   538 AA;  59513 MW;  1189AD6C5C726996 CRC64;

  Query Match             73.7%;  Score 2041;  DB 10;  Length 538;
  Best Local Similarity   73.8%;  Pred. No. 3.1e-148;
  Matches  399;  Conservative   57;  Mismatches   71;  Indels   14;  Gaps    4;

Qy      5 LSSLLQN-SDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLCTGV 63
          | ||:     | ||||||||::||||||||||||||||||||:|||||||:||| |||
Db      2 LDSLVSKLPSLSTSDHASVVALNLFVALLCACIVLGHLLEENRWMNESITALLIGLGTGV 61

Qy     64 VILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLIS 123
             |||:| |||||||||||||||||||||||||||||||||||| ||:|||||||:||:||
Db     62 TILLISKGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFRNFVTIMLFGAVGTIIS 121

Qy    124 CSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVFGEG 183
          |:||| |  : || |||   | |||||||||||||||||||:|||||||||||||||
```

```
Db    122 CTIISLGVTQFFKKLDIGTFDLGDYLAIGAIFAATDSVCTLQVLNQDETPLLYSLVFGEG 181

Qy    184 VVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKLY 243
          |||||||||:||||||||:|  : :   | :|||||||| || ||   ||: ||:|||||
Db    182 VVNDATSVVVFNAIQSFDLTHLNHEAAFHLLGNFLYLFLLSTLLGAATGLISAYVIKKLY 241

Qy    244 FGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRH 303
          |||||||||||||||:||||::|||| |||||||||||||||||||||||||||:||:|
Db    242 FGRHSTDREVALMMLMAYLSYMLAELFDLSGILTVFFCGIVMSHYTWHNVTESSRITTKH 301

Qy    304 SFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLS 363
          :|||||:|||||||||||||||:||: | :: | |:|||||:||::||||||||||
Db    302 TFATLSFLAETFIFLYVGMDALDIDKWRSVSDTPGTSIAVSSILMGLVMVGRAAFVFPLS 361

Qy    364 FLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTV 423
          |||||||| |:||:|  |::|||:||||||||:|||||||| :||| :  ||||||||:
Db    362 FLSNLAKKNQSEKINFNMQVVIWWSGLMRGAVSMALAYNKFTRAGHTDVRGNAIMITSTI 421

Qy    424 TVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTT---SEPSSPKHFTVPLLDNQPDS 480
          || |||||||::|||||: |||        | : ||   |: ::||   :||||   ||
Db    422 TVCLFSTVVFGMLTKPLISYLLP--------HQNATTSMLSDDNTPKSIHIPLLDQ--DS 471

Qy    481 ESDMITGPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVEQS 540
            :      | || ::|  |  || ||| |||:|||||||||||||||||| ||| |::
Db    472 FIEPSGNHNVPRPDSIRGFLTRPTRTVHYYWRQFDDSFMRPVFGGRGFVPFVPGSPTERN 531

Qy    541 P 541
          |
Db    532 P 532


RESULT 11
Q94LX3
ID   Q94LX3      PRELIMINARY;      PRT;   555 AA.
AC   Q94LX3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NA H-ANTIPORTOR.
GN   THPURPLE.
OS   Torenia hybrida.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Lamiales; Lamiales incertae sedis; Torenieae;
OC   Torenia.
OX   NCBI_TaxID=75807;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Iida S., Kusumi T., Yonekura-Sakakibara K., Tanaka Y.;
RT   "Plant Na-H antiportor.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB051819; BAB56107.1; -.
SQ   SEQUENCE   555 AA;  61315 MW;  FEF9556E029B3983 CRC64;

  Query Match             70.6%;  Score 1953.5;  DB 10;  Length 555;
  Best Local Similarity   70.5%;  Pred. No. 1.6e-141;
  Matches  389;  Conservative   56;  Mismatches   88;  Indels   19;  Gaps    4;

Qy      4 GLSSLL-----QNSDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIG 58
          |  |::     :  :|::| | |||:: ||| ||| |||:|||||||||:|||| |||||
Db      2 GFESVIKLAASETDNLWSSGHGSVVAITLFVTLLCTCIVIGHLLEENRWMNESIIALIIG 61

Qy     59 LCTGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAI 118
          | |||:|||:|||||||||||||||||||| |||||||||||||| || ||:|||:
Db     62 LATGVIILLISGGKSSHLLVFSEDLFFIYALPPIIFNAGFQVKKKSFFRNFATIMMFGAV 121

Qy    119 GTLISCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSL 178
          |||||  ||| | :  |  :::  |  |||||||||||||||||||||||||||||||||
Db    122 GTLISFIIISLGTIAFFPKMNMR-LGVGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSL 180

Qy    179 VFGEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYI 238
          |||||||||||||||||:|:||:           :||| ||| :|| |||  ||| |||
Db    181 VFGEGVVNDATSVVLFNAVQNFDLPHMSTAKAFELVGNFFYLFATSTVLGVLTGLLSAYI 240
```

```
Qy    239 IKKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSR 298
          ||||||||||||||||:|:||:||||::|||| ||||||||||||||||||||||||:|| 
Db    241 IKKLYFGRHSTDREVAIMILMAYLSYMLAELFDLSGILTVFFCGIVMSHYTWHNVTENSR 300

Qy    299 VTTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAF 358
          |||:|:|||||||| ||||||||||||||:|| |   | |||: |:||:|: ||||
Db    301 VTTKHTFATLSFVAEIFIFLYVGMDALDIEKWRFVSGSMTTSAAVSATLLGLVLLSRAAF 360

Qy    359 VFPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIM 418
          |||||||||||||: :||| |||||||||||||||:|||| :||  | |   ||||
Db    361 VFPLSFLSNLAKKSPLEKISLRQQIIIWWAGLMRGAVSMALAYKQFTREGLTVERENAIF 420

Qy    419 ITSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQP 478
          ||||:|:|||||||||||||||||:|  |   |      :||| :|   |:|||
Db    421 ITSTITIVLFSTVVFGLMTKPLINLLIPSPKLNRS-----VSSEPLTPNSITIPLLGESQ 475

Qy    479 DSESDMIT--------GPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVP 530
          || ::: :        |  ||||::|||||  |||||| ||||||::|||||||||||||
Db    476 DSVAELFSIRGQTSQGGEPVARPSSLRMLLTKPTHTVHYYWRKFDNAFMRPVFGGRGFVP 535

Qy    531 FVAGSPVEQSPR 542
          :| ||| |:| |
Db    536 YVPGSPTERSVR 547


RESULT 12
Q9SQU0
ID   Q9SQU0      PRELIMINARY;      PRT;   552 AA.
AC   Q9SQU0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE SODIUM PROTON EXCHANGER.
GN   F24P17.16.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
RA   Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA   Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT   "Arabidopsis thaliana chromosome III BAC F24P17 genomic sequence.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC011623; AAF08577.1; -.
DR   InterPro; IPR001179; FKBP_PPIase.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
DR   PROSITE; PS00453; FKBP_PPIASE_1; UNKNOWN_1.
SQ   SEQUENCE   552 AA;  61135 MW;  FB4317D8A874FCE9 CRC64;

  Query Match           69.6%;  Score 1925.5;  DB 10;  Length 552;
  Best Local Similarity   70.4%;  Pred. No. 2.2e-139;
  Matches  381;  Conservative   59;  Mismatches   94;  Indels    7;  Gaps    3;

Qy      1 MAFGLSSLLQNSD-LFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGL 59
          |  |||::|: :: || ||||||||||||||||||||||||||||| ||:||||||||||
Db      1 MVIGLSTMLEKTEALFASDHASVVSMNLFVALLCACIVLGHLLEETRWMNESITALIIGS 60

Qy     60 CTGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIG 119
          |||:||||:|||||| :|||||||||||||||||||||||||||| ||||||||||||
Db     61 CTGIVILLISGGKSSRILVFSEDLFFIYLLPPIIFNAGFQVKKKQFFRNFMTIMLFGAIG 120

Qy    120 TLISCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLV 179
          ||||  ||||||  :|: ::|  |   |||||||||:||||||||||:||||||||||||
Db    121 TLISFVIISFGAKHLFEKMNIGDLTIADYLAIGAIFSATDSVCTLQVLNQDETPLLYSLV 180

Qy    180 FGEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYII 239
          |||||||||||||||||| ||:|: :  | | | ||| |||: || |||  ||| |::|
Db    181 FGEGVVNDATSVVLFNAIQRFDLTNINSAIALEFAGNFFYLFILSTALGVAAGLLSAFVI 240

Qy    240 KKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRV 299
          |||| ||||||||||||::||||::||||:|| |||||||||||||||||||||||: |:|
Db    241 KKLYIGRHSTDREVALMMLLAYLSYMLAELFHLSSILTVFFCGIVMSHYTWHNVTDKSKV 300

Qy    300 TTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFV 359
          ||:|:|| :||:|| |||||||||||||||| |:|| | |: |||||:||||:||||||
Db    301 TTKHTFAAMSFLAEIFIFLYVGMDALDIEKWDVVRNSPGQSIGVSSILLGLILLGRAAFV 360

Qy    360 FPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMI 419
          |||||||| | :  :||  ::|: ||||||||||||:|||||:|||||| :  |||||
Db    361 FPLSFLSNLTKSSPDEKIDLKKQVTIWWAGLMRGAVSMALAYNQFTTSGHTKVLGNAIMI 420

Qy    420 TSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPD 479
          |||:|||||||||||:||||:  | |  || :   :|              |||  |
Db    421 TSTITVVLFSTVVFGLLTKPLVKHLQPSSKQSSTTALQITLRSSFHDPILHEPLLSTQGQ 480

Qy    480 SESDMITGPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVEQ 539
          || |    ||  :  : ||  ::|:  :| ||||||:: || :|||||  | | |||:|
Db    481 SEYD----PE--QHVSFRMFWKSPSRAIHHYWRKFDNAVMRRIFGGRGVSPVVPGSPIEN 534

Qy    540 S 540
          |
Db    535 S 535


RESULT 13
Q94BM4
ID   Q94BM4      PRELIMINARY;      PRT;   546 AA.
AC   Q94BM4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NA+/H+ ANTIPORTER.
GN   NHX1.
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Triticum.
OX   NCBI_TaxID=4565;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wang Z., Zhang J., Chen S.;
RT   "Isolation and characterization of two Na+/H+ antiport genes from
RT   wheat.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY040245; AAK76737.1; -.
SQ   SEQUENCE   546 AA;  59704 MW;  45E1BCAB73E295C6 CRC64;

  Query Match           67.4%;  Score 1865.5;  DB 10;  Length 546;
  Best Local Similarity   68.2%;  Pred. No. 8.7e-135;
  Matches  369;  Conservative   67;  Mismatches   92;  Indels   13;  Gaps    4;

Qy      1 MAFGLSSL-LQNSDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGL 59
          |   | :| |: : |  ||| |:|::|:|:|||| ||| ||||| ||||||| |||::||
Db      1 MGLDLGALALKYTGLAVSDHDSIVAINIFIALLCGCIVFGHLLEGNRWVNESTTALVLGL 60

Qy     60 CTGVVILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIG 119
           || |||: : | :| :|:|||||:|||||||||||||||||||||| || ||:|||| |
Db     61 ITGGVILICTKGVNSRILIFSEDIFFIYLLPPIIFNAGFQVKKKQFFRNFATIILFGAAG 120

Qy    120 TLISCSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLV 179
          ||||  ||:|||: :|  ||:  |: |||||||||:|||||||||||:||| |||||||
Db    121 TLISFVIITFGAMGLFSKLDVGPLELGDYLAIGAIFSATDSVCTLQVLNQDEAPLLYSLV 180

Qy    180 FGEGVVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYII 239
          |||||||||||||||||: |:  ||  : | ||| ||||| :|| |||  ||| ||||
Db    181 FGEGVVNDATSVVLFNAIQNIDINHFDVFVLLQFIGKFLYLFFTSTVLGVAAGLLSAYII 240

Qy    240 KKLYFGRHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRV 299
          ||| | |||||||||:|:||:||||::: |  ||||||||||||||||||||||||||||
```

```
Db    241 KKLCFARHSTDREVAIMILMAYLSYMLSMLLDLSGILTVFFCGIVMSHYTWHNVTESSRV 300

Qy    300 TTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFV 359
          ||:|:||||||:|| |:|||||||||||:|||   :|    :|:|::::||::|||||||
Db    301 TTKHTFATLSFIAEIFLFLYVGMDALDIDKWKLASSSPKKPIALSAVILGLVMVGRAAFV 360

Qy    360 FPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMI 419
          |||||||||:|| |  |||| ||:|||||||||||||||||||||||::  ||:||
Db    361 FPLSFLSNLSKKESHPKISFNQQVIIWWAGLMRGAVSIALAYNKFTTSGHTAVRVNAVMI 420

Qy    420 TSTVTVVLFSTVVFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPD 479
          |||: ||||||:||||:||||||||:||     :  || :  :|      |  ||
Db    421 TSTIIVVLFSTMVFGLLTKPLINLLIPPRPGTAADISSQSFLDP-----LTASLLG---- 471

Qy    480 SESDMITGPEVARPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVPFVAGSPVEQ 539
            ||   | ::   | |: ||  || :||| |||||| ||||:|||||||||| |||:|:
Db    472 --SDFDVG-QLTPQTNLQYLLTMPTRSVHRVWRKFDDKFMRPMFGGRGFVPFVPGSPIER 528

Qy    540 S 540
          |
Db    529 S 529
```

```
RESULT 14
Q93YH2
ID   Q93YH2      PRELIMINARY;      PRT;   534 AA.
AC   Q93YH2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NA+/H+ ANTIPORTER, ISOFORM 1.
GN   NHX1.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Venema K., Belver A., Donaire J.;
RT   "Cloning and characterization of two Na+/H+ antiporter isoforms from
RT   tomato.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ306630; CAC84522.1; -.
SQ   SEQUENCE   534 AA;  59037 MW;  E4E736FA4A0C0791 CRC64;

  Query Match             65.7%;  Score 1819.5;  DB 10;  Length 534;
  Best Local Similarity   68.6%;  Pred. No. 2.8e-131;
  Matches  356;  Conservative   66;  Mismatches   88;  Indels    9;  Gaps    3;

Qy     12 SDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLCTGVVILLLSGG 71
          | |   |  || |: ||||||| |||:||||||:||:|:||| |:||| || :||| : |
Db     12 SILSDGDQVSVDSITLFVALLCGCIVIGHLLEESRWINDSITTLVIGLSTGGIILLTTKG 71

Qy     72 KSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLISCSIISFGA 131
          |||||| | | |||||:||||||||||||||||| ||:|||||||:||||| ||||||
Db     72 KSSHLLEFDEQLFFIYVLPPIIFNAGFQVKKKQFFRNFVTIMLFGAVGTLISFSIISFGA 131

Qy    132 VKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVFGEGVVNDATSV 191
           ::   ||| ||:  |||||||||:|||||||| |:||||| ||||||||||||||||||
Db    132 KELLGKLDIGFLELRDYLAIGAIFSATDSVCTLQALNQDETPRLYSLVFGEGVVNDATSV 191

Qy    192 VLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGIGLLCAYIIKKLYFGRHSTDR 251
          |||||||  |::  : :  | | |||||||||:|||||| |||| ||:|||:| |||||||
Db    192 VLFNAIQKLDLSHINSRAALVFTGNFLYLFLASTFLGVLIGLLSAYLIKKIYLGRHSTDR 251

Qy    252 EVALMMLMSYLSYIMAELFYLSGILTVFFCGIVMSHYTWHNVTESSRVTTRHSFATLSFV 311
          |||||:||:||||:||||| |||||||| |||||||||||||| :|:||||:|||||||:
Db    252 EVALMILMAYLSYVMAELFDLSGILTVFICGIVMSHYTWHNVTFNSKVTTRHAFATLSFI 311

Qy    312 AETFIFLYVGMDALDIEKWKFVKNSQGLSVAVSSILVGLILVGRAAFVFPLSFLSNLAKK 371
          || |||||||||||||||:|||:| | || ||: |:||:||||| ||||||  ||  |:
Db    312 AEIFIFLYVGMDALDIEKWRFVKDSPGKSVGVSAALLGLVLVGRACFVFPLSLFSNCLKR 371
```

```
Qy    372 NSSDKISFRQQIIIWWAGLMRGAVSIALAYNKFTTSGHTSLHENAIMITSTVTVVLFSTV 431
          :  ||   : |: |||||||||:||:|||||:||  |||    ||:|||||:|:||||||
Db    372 SEHDKFGLKLQVTIWWAGLMRGSVSMALAYNQFTRFGHTQQPGNAVMITSTITIVLFSTV 431

Qy    432 VFGLMTKPLINLLLPPHKQMPSGHSSMTTSEPSSPKHFTVPLLDNQPDSESDMITGPEVA 491
          ||||:|||||:  |||  :    | ::: :|| |    |  ||| |: : | :|    :
Db    432 VFGLITKPLVRFLLPSSQ----GFNNLISSEQS----FARPLLTNEQELELEM-GNVDPV 482

Qy    492 RPTALRMLLRTPTHTVHRYWRKFDDSFMRPVFGGRGFVP 530
          ||: | :||: |::|:| :||:|||:||||:||||||||
Db    483 RPSGLSILLKEPSYTIHNHWRRFDDAFMRPLFGGRGFVP 521
```

```
RESULT 15
O04655
ID   O04655      PRELIMINARY;      PRT;   457 AA.
AC   O04655;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   A_TM021B04.4 PROTEIN.
GN   A_TM021B04.4.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Dante M., Wamsley P., Gibson A.;
RT   "The sequence of A. thaliana TM021B04.";
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Wash-U;
RT   "The A. thaliana Genome Sequencing Project.";
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Waterston R.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF007271; AAB61069.1; -.
DR   InterPro; IPR001179; FKBP_PPIase.
DR   InterPro; IPR000676; NaH_Exchngr.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PROSITE; PS00453; FKBP_PPIASE_1; UNKNOWN_1.
SQ   SEQUENCE   457 AA;  50611 MW;  0AF2F235F1A258EE CRC64;

  Query Match             62.6%;  Score 1732.5;  DB 10;  Length 457;
  Best Local Similarity   75.4%;  Pred. No. 1.1e-124;
  Matches  341;  Conservative   44;  Mismatches   46;  Indels   21;  Gaps    3;

Qy      5 LSSLLQN-SDLFTSDHASVVSMNLFVALLCACIVLGHLLEENRWVNESITALIIGLCTGV 63
          | ||:     | |||||||::||||||||||||||||||||:|||||||:||| |||
Db      2 LDSLVSKLPSLSTSDHASVVALNLFVALLCACIVLGHLLEENRWMNESITALLIGLGTGV 61

Qy     64 VILLLSGGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFVNFMTIMLFGAIGTLIS 123
            |||:| |||||||||||||||||||||||||||||||||| ||:|||||||:||:||
Db     62 TILLISKGKSSHLLVFSEDLFFIYLLPPIIFNAGFQVKKKQFFRNFVTIMLFGAVGTIIS 121

Qy    124 CSIISFGAVKIFKHLDIDFLDFGDYLAIGAIFAATDSVCTLQVLSQDETPLLYSLVFGEG 183
          |:||| |  : || |||   | |||||||||||||||||||:|||||||||||||||
Db    122 CTIISLGVTQFFKKLDIGTFDLGDYLAIGAIFAATDSVCTLQVLNQDETPLLYSLVFGEG 181

Qy    184 VVNDATSVVLFNAIQSFDMTSFDPKIGLHFIGNFLYLFLSSTFLGVGI------------ 231
          |||||||||:|||||||||:|  : :   | :|||||||| || ||  :
Db    182 VVNDATSVVVFNAIQSFDLTHLNHEAAFHLLGNFLYLFLLSTLLGAAVSPLFSSLPFFLT 241

Qy    232 GLLCAYIIKKLYFG--------RHSTDREVALMMLMSYLSYIMAELFYLSGILTVFFCGI 283
```

```
          ||: ||:|||||||        |||||||||||||||:||||::|||| ||||||||||||
Db    242 GLISAYVIKKLYFGRWPHINCHRHSTDREVALMMLMAYLSYMLAELFDLSGILTVFFCGI 301

Qy    284 VMSHYTWHNVTESSRVTTRHSFATLSFVAETFIFLYVGMDALDIEKWKFVKNSQGLSVAV 343
          |||||||||||||||:||:|:||||||:||||||||||||||||:||: | :: | |:||
Db    302 VMSHYTWHNVTESSRITTKHTFATLSFLAETFIFLYVGMDALDIDKWRSVSDTPGTSIAV 361

Qy    344 SSILVGLILVGRAAFVFPLSFLSNLAKKNSSDKISFRQQIIIWWAGLMRGAVSIALAYNK 403
          ||||:||::|||||||||||||||||||| |:||:|  |::|||:||||||||:||||||
Db    362 SSILMGLVMVGRAAFVFPLSFLSNLAKKNQSEKINFNMQVVIWWSGLMRGAVSMALAYNK 421

Qy    404 FTTSGHTSLHENAIMITSTVTVVLFSTVVFGL 435
          || :||| :  ||||||||:|| ||||||  |
Db    422 FTRAGHTDVRGNAIMITSTITVCLFSTVVRNL 453

Search completed: October 18, 2002, 12:26:45
Job time : 34 secs
```